



**PRONOMINAL SUBJECTS IN THE ENGLISH OF ARABIC,
FINNISH AND FRENCH SPEAKERS**

Naif Saeed Alsaedi

Thesis submitted to the School of English Literature, Language and
Linguistics for the Degree of

Doctor of Philosophy (PhD in Linguistics)

at

Newcastle University

July 2017

Abstract

Previous studies designed to investigate whether null-subject parameter settings transfer in second-language acquisition (L2A/SLA) have produced inconclusive, differing, and even conflicting results. While some researchers claim that the first language (L1) value of the parameter does not transfer into L2A, others argue that it does; furthermore, they disagree about whether its L1 value could be reset to a value appropriate to the second language (L2) (i.e., White, 1985; Hilles, 1986; Phinney, 1987; Tsimpli and Roussou, 1991; Al-Kasey and Pérez-Leroux, 1998; Liceras and Díaz, 1999; LaFond, 2001; Sauter, 2002; Judy, 2011; Orfitelli and Grüter, 2013). The aim of this study is to address these issues in a more accurate way by paying attention to a number of factors both internal and external to learners that have been overlooked in previous studies, resulting in conflicting conclusions about null-subject transfer and parameter resetting in L2A. This study investigates the acquisition of the obligatory overt subject pronouns in English by three groups of learners whose L1s belongs to three distinct language types – namely, non-null subject languages (French), partial-null subject languages (Finnish), and consistent-null subject languages (Arabic). The participants in each group were divided into three subgroups – lower intermediate, upper-intermediate, and advanced – on the basis of their scores on the proficiency test in order to examine how the investigated L2 grammar changes at the different developmental stages in relation to the learners' different native languages. The data were collected from 487 participants by means of a grammaticality judgement (GJ) task and a translation task. The findings from the GJ task show evidence that all learners, regardless of linguistic background, start off with pro-drop and then transfer their L1 parameter setting at the intermediate and late stages of L2A, whereas the findings from the translation task suggest that the L1 setting of the null subject parameter transfers in L2A. However, the results show that there are structural, developmental, and situational/contextual (realised as task-type) constraints on when, where, and to what extent pronominal subjects can be null. The results indicate that learners persistently accept referential embedded null subjects in the GJ task beyond the stage of L2 development when they have established the requirement for overt subjects in their production. Moreover, the results provide evidence that all participants, as proficiency subgroups, regardless of their L1 backgrounds, treated null subjects in the two types of experimental sentences differently; they accepted significantly fewer null subjects in complement clauses than in adverbial clauses. However, only the French participants converged on the target grammar in all respects; the Arabic and the Finnish participants continued to perform non-target-like like in their judgement of null subjects, if only in adverbial clauses. Group results indicating that L2 learners' performance varies from task to task and from structure to structure suggest that null subject parameter settings cannot be reset in L2A. These findings, which show that there are structural and situational or contextual constraints on when and where pronominal subjects can be null, suggest that L2 learners rely on discourse licensing of null subjects. In other words, the results indicate that argument omission vs. overt expression in L2 depends on the referent's discourse status, which can be defined in terms of a range of discourse and pragmatic notions. The results also raise and leave unanswered several questions that require further investigation.

**This work is dedicated, with all my love, to my family and the loving
memory of my grandfather.**

Acknowledgments

First and foremost, I wish to express my deepest gratitude to my thesis supervisors, Professor Martha Young-Scholten and Professor Anders Holmberg, for their invaluable feedback, advice, and guidance throughout my time in the PhD program. I am very fortunate to have had them as my supervisors. I am also grateful to the Newcastle University linguistics staff, Noel Burton-Roberts, Heike Pichler, Hermann Moisl, and Joel Wallenberg, for having taken the time to discuss my work with me. My gratitude is also due to Professor Roger Hawkins at Essex University, who first taught me second-language acquisition. Special thanks go to all of my research participants at the University of Helsinki, Aix Marseille University, Bordeaux University, Taibah University, and Umm Al-Qura University. In this regard, I must especially thank Inna Smirnova, Emanuela Costea, Bibi Walker, Joachim Zemmour, Dr Ali Jarah, Ahmed Alhassani, Nadia Alsaedi and Afaf Alsaedi, who helped me recruit the participants in my study. Without a doubt, their effort trebled the number of participants and greatly improved the thesis. I must also acknowledge the invaluable help I received from Omar Saeed and Majid Alsaedi with statistical analyses of the data. I would like to take this opportunity to thank my internal and external examiners Dr Nick Riches at Newcastle and Dr Anne Vainikka at the University of Delaware; this thesis and I have greatly benefited from their insightful comments. I would also like to express my deepest appreciation to my parents, brothers, and sisters for their constant love, continuous support, and endless encouragement. Most of all, though, I should like to thank my closest friend—my wife, Nada. Her unswerving love, unwavering support, encouragement, and patience made it possible for me to keep striving to reach my goal. Without you, Nada, I would not have been able to discover the joy of life during those difficult days; I will be grateful to you forever. Last but not the least, I gratefully acknowledge the financial support I have received from Taibah University during my time at Newcastle University.

Table of Contents

Abstract	i
Acknowledgment	iii
Table of Contents	v
List of Tables	x
List of Figures	xiv
List of Abbreviations	xvi
Chapter 1. General Introduction	1
Chapter 2. UG Theory and Language Acquisition: Evidence from the Null Subject Parameter	4
2.1 Overview of the Chapter	4
2.2 UG and Child Language Acquisition	5
2.3 The Nature of the Language Faculty	8
2.3.1 Principles and Parameters Theory and the Minimalist Program for Linguistic Theory	8
2.3.2 The Null Subject Parameter(s)	10
2.3.3 Similarities and differences among Arabic and Finnish	13
2.3.4 The syntax of the null-subject phenomenon: licensing and identification within the agreement-based analysis	17
2.4 Null subjects in first language acquisition	20
2.5 The role of generative grammar in the study of first language acquisition: concluding remarks	25

Chapter 3. The Role of UG and the Status of L1 Null-Subject Parameter-Setting Transfer in L2 Acquisition	26
3.1 Introduction	26
3.2 Universal Grammar and Second-Language Acquisition	26
3.3 Review of L2 Null Subject Parameter Studies.....	37
3.3.1 Early Studies.....	39
3.3.2 Recent studies.....	45
Chapter 4. Rationale and Experimental Design.....	49
4.1 Introduction	49
4.2 Research Gap, Questions, and Hypotheses	50
4.2.1 Research Gap and Motivation for the Study	50
4.2.2 Research Questions and Sub-questions	52
4.2.3 Research Hypotheses	52
4.3 Methodology	55
4.3.1 Participants.....	55
4.3.2 Test Instruments: Their Validity, Reliability, and Design.....	61
4.3.2.1 The Grammaticality Judgement Task.....	62
4.3.2.1.1 The Validity and Reliability of the GJ Test	62
4.3.2.1.2 The Grammaticality Judgement Test Design.....	65
4.3.2.2 The Translation Test.....	83
4.3.2.2.1 The Validity and Reliability of the Translation Test.....	83
4.3.2.2.2 The Translation Test Design	85
4.3.3 Piloting the Instrument.....	89
4.3.3.1 Initial Piloting.....	89
4.3.3.2 Final Piloting	92
4.3.4 Data-Collection Procedures	98

4.3.4.1	The Consent Form	98
4.3.4.2	The Personal Information Sheet	100
4.3.4.3	The Administration of the Data Elicitation Tools	100
4.3.4.4	The Debriefing Form	102
4.3.5	Measuring the L2 Participants' Levels of Language Proficiency: Challenge and Solution	102
4.3.6	Scoring: Procedure and Method.....	109
4.3.6.1	The GJ Task: The Adopted Marking Method.....	109
4.3.6.2	The Translation Task: The Adopted Marking Method.....	116
4.3.7	Data Analytical and Statistical Framework	119
Chapter 5.	Results and Discussion	124
5.1	Introduction.....	124
5.2	Results and discussion by task	125
5.2.1	The grammaticality judgment task.....	126
5.2.1.1	The English ILs of the French, Finnish, and Arabic speakers: Basic descriptive and inferential statistics	126
5.2.1.2	Comparisons across all language subgroups: ILs vs. ILs of the same proficiency level	130
5.2.1.3	Discussion of grammaticality judgement tasks' results.....	133
5.2.2	The translation task.....	143
5.2.2.1	The English IL of the French, Finnish and Arabic speakers: Basic descriptive and inferential statistics	143
5.2.2.2	Comparisons across all language subgroups: ILs vs. ILs of the same proficiency level	147
5.2.2.3	Discussion of the translation task's results	150
5.2.3	The grammatical intuition data vs. translational production data: Comparisons within each subgroup	152

5.2.3.1 Comparisons on subjects' null realisation across the two tasks	153
5.2.3.2 Discussion of results divergence/variability across the two tasks	156
5.3 Results and discussion by grammatical constructions	163
5.3.1 The translation task: Subject drop in complement clauses vs. in adverbial clauses	164
5.3.1.1 The English IL of the French, Finnish and Arabic speakers compared within subgroups.....	164
5.3.1.2 Discussion of the translation task's results by grammatical structures	167
5.3.2 The Grammaticality Judgment Task: Null Subject Acceptance in Complement Clauses vs. Adverbial Clauses	169
5.3.2.1 The English IL of the French, Finnish, and Arabic speakers: Compared within subgroups	169
5.3.2.2 Discussion of the GJ task's results by grammatical structures ..	176
5.4 Referential null subjects licensing in SLA.....	185
Chapter 6. Conclusion.....	199
6.1 Overview of the Chapter	199
6.2 Summary of the Study	199
6.3 Limitations of the Study	203
6.4 Suggestions for Further Research.....	205
Appendices	207
Appendix 1: Tests instruments.....	208
Appendix 1.a. The GJ task and the translation task (French version)	202
Appendix 1.b. The GJ task and the translation task (Finnish version)	220
Appendix 1.c. The GJ task and the translation task (Arabic version)	229
Appendix 2: Lists of English words with their meanings in the source languages	244

Appendix 2.a. French version.....	238
Appendix 2.b. Finnish version.....	239
Appendix 2.c. Arabic version.....	240
Appendix 3. Initial pilot experimental test version	247
Appendix 4. Consent forms	259
Appendix 4.a. Version for the adult participants	253
Appendix 4.b. Version for the minor participants	255
Appendix 5. Personal information form	263
Appendix 6. Debriefing form	264
Appendix 7. Oxford Online Placement Test	265
Appendix 8. List of Shapiro Results.....	273
Appendix 9. Descriptive statistics: acceptance of null subjects with local antecedents vs. with non-local antecedents	277
Appendix 10. Proposed descriptive statistics for the English native-speakers control group showing number of dropping subjects	279
Appendix 11. Descriptive statistics: distributions of items with 3 rd person singular null subjects and items with missing subject-verb agreement	2798
11-a. Grammaticality judgment task	288
11-b. The translation task	289
Appendix 12. Descriptive statistics: distributions of items with 3 rd person null subjects and items with 1 st or 2 nd person null subjects	283
12-a. Grammaticality judgment task	276
12-b. The translation task	277
Appendix 13. Further statistical analysis and discussion relevant to footnote number 100	286
References	288

List of Tables

Table 2-1. Subject-verb agreement morphology of the present tense of the verb puhu ‘speak’ and its Arabic counterpart.....	18
Table 2-2. Percentages of Null Subjects across Child Languages	22
Table 3-1. Subject types used in the production task of Orfitelli and Grüter (2013).	48
Table 4-1. Participants involved in the study.....	55
Table 4-2. Participants included in the study.....	61
Table 4-3. Oxford Online Placement Test: proficiency-levels classifications scale	106
Table 4-4. Reclassification of Oxford Online Placement Test proficiency-levels scale	106
Table 4-5. Assessing accuracy: percentage of uncertainty associated with levels of proficiency: a comparison between the original and the reclassified scales	107
Table 4-6. Oxford Online Placement Test reclassified levels of proficiency: score range redistributing/remapping	108
Table 4-7. Methods used to mark the GJ test.....	113
Table 4-8. Classifying the possible responses to the judged sentences: included responses vs. excluded responses.....	113

Table 4-9. Calculation processes used to exclude reactions and participants in the GJ task.....	114
Table 4-10. Method used in regrouping the included reactions: accepted vs. rejected English sentences.....	114
Table 4-11. Possible translational performances: included performances vs. excluded performances	117
Table 4-12. Method used to mark the translation test	118
Table 4-13. Calculation procedures used to exclude performances and participants in the translation task	119
Table 5-1. Descriptive statistics for French-speaking learners' acceptance of ungrammatical sentences with null embedded subjects	127
Table 5-2. Descriptive statistics for the English native-speakers control group: acceptance of ungrammatical sentences with null embedded subjects	127
Table 5-3. Descriptive statistics for L1 Finnish- and Arabic-speaking learners' acceptance of ungrammatical sentences with null embedded subjects	128
Table 5-4. Inter-subgroup inferential comparisons between the L2 learners (L1 Finnish and L1 Arabic) and the native-English controls	129
Table 5-5. The results of the statistical tests comparing the different subgroups of the participants on acceptance of null subjects.....	132
Table 5-6. White's results (1985): French responses by level to individual sentences with missing subjects: number responding "correct." to ungrammatical sentences.....	134
Table 5-7. White's results (1985): French responses by level to all sentences with missing subjects: percentage of acceptance	135

Table 5-8. Inferential comparisons between the Finnish learners and the French learners on embedded null subjects with local antecedents and with non-local antecedents.....	139
Table 5-9. Descriptive statistics for ungrammatical sentences with missing subjects produced by French-speaking learners.....	144
Table 5-10. Descriptive statistics for the ungrammatical sentences with null subjects produced by Finnish-speaking learners.....	145
Table 5-11. Descriptive statistics for ungrammatical sentences with null subjects produced by Arabic participants.....	147
Table 5-12. Summary table of mean scores of sentences with null subjects produced by French, Finnish, and Arabic L2 learners of English.....	164
Table 5-13. Descriptive statistics by items types (complement vs. adverbial) for sentences with null subjects produced by lower-intermediate subgroups of learners.....	165
Table 5-14. Inferential comparisons between the lower-intermediate French and Finnish participants and the native controls: subject drop in complement clauses vs. in adverbial clauses.....	167
Table 5-15. L1 French acceptances by level of sentences with null subjects in embedded complement and adverbial clauses.....	170
Table 5-16. Comparisons between the French participants and the native English controls: acceptance of null subjects in complement vs. in adverbial clauses.....	172
Table 5-17. L1 Finnish acceptances by level to sentences with null subjects in embedded complement and adverbial clauses.....	172

Table 5-18. Comparisons between the Finnish participants and the native English controls: acceptance of null subjects in complement vs. in adverbial clauses	174
Table 5-19. L1 Arabic participants acceptances by level of sentences with null subjects in embedded complement and adverbial clauses.....	174
Table 5-20. Comparisons between the Arabic-speaking participants and the native English controls: acceptance of null subjects in complement vs. adverbial clauses.....	176
Table 5-21. Correlations between third-person singular null subjects and missing subject-verb agreements.....	189
Table 5-22. Inferential comparisons between acceptance of third-person singular null subjects and acceptance of missing subject-verb agreements	190
Table 5-23. Statistical intra-subgroup comparisons between acceptance of third-person null subjects and acceptance of first- and second-person null subjects	194
Table 5-24. The relationship between L1 and acceptance of 3rd-person null subjects and 1st- or 2nd-person null subjects	199

List of Figures

Figure 2-1. Model of child L1 acquisition	21
Figure 4-1. The participants' shared and variable characteristics	60
Figure 5-1. Summary and comparison of the subgroups' results in terms of means of null subject acceptance.....	131
Figure 5-2. Percentages of null subject acceptances with local antecedents vs. non-local antecedents.	138
Figure 5-3. Percentages of null subject acceptances with local antecedents vs. non-local antecedents.	141
Figure 5-4. Intergroup comparisons between the subgroups of learners of the same proficiency level.....	148
Figure 5-5. Pairwise comparisons between the translation task and the GJ task performances within each lower-intermediate-level subgroup of learners.	153
Figure 5-6. Pairwise comparisons between the translation task and the GJ task performances within each upper-intermediate-level subgroup of learners.	154
Figure 5-7. Pairwise comparisons between the translation task and the GJ task performances within each advanced subgroup of learners.....	155
Figure 5-8. The subgroups' overall percentages of the PIT scores to ungrammatical sentences with null subjects.	158
Figure 5-9. Number of sentences with null subjects accepted by every individual Finnish advanced participant.	162
Figure 5-10. Percentages of subject drop in complement clauses and in adverbial clauses by the French, Finnish, and Arabic lower-intermediate learners of English.	169

Figure 5-11. L1 French acceptance by level of sentences with null subjects in embedded complement and adverbial clauses.	171
Figure 5-12. L1 Finnish acceptances by level of sentences with null subjects in embedded complement and adverbial clauses.	173
Figure 5-13. L1 Arabic-speaking participants' acceptances by level of sentences with null subjects in embedded complement and adverbial clauses.	175
Figure 5-14. Number of embedded adverbial clauses with null subjects accepted by Finnish advanced-level learners of English.	183
Figure 5-15. Participants' IL systems: Percentage of acceptance of items with third-person singular null subjects and items with missing subject-verb agreement.	187
Figure 5-16. Learners' performance on the translation task: Percentages of third-person subject drops vs. percentages of missing subject-verb agreement.	191
Figure 5-17. Intra-subgroup comparison of non-target performance in the acceptance of null subjects with third-person referents and with first- or second-person referents in the GJ task.	194

List of Abbreviations

ADV	Advanced
AGR	Agreement
CA	Contrastive Analysis
CAH	Contrastive Analysis Hypothesis
CIA	Contrastive Interlanguage Analysis
CPH	Critical Period Hypothesis
D	Definite
DP	Determiner Phrase
EFL	English as a Foreign Language
ESL	English as a Second Language
FLB	Faculty of language in narrow sense
FLN	Faculty of language in broad sense
FT/FA	Full Transfer/Full Access Hypothesis
FUT	Future
G ₀	The Initial State of Language Faculty
GJ task	Grammaticality judgment task
GJs	Grammaticality judgments
GLMM	Generalised linear mixed models
H1	Hypothesis 1
H2	Hypothesis 2
H3	Hypothesis 3
H4	Hypothesis 4
IL	inter-language
I-language	The Internalized Linguistic System
L1	First Language
L2	Second Language

L2A	The Process of Second Language Acquisition
L3	Third language
L3A	Third Language Acquisition
LI	lower-intermediate
LP	Language Proficiency
Max	Maximum
Min	Minimum
N	Number
NOM	Nominative
Non-NSL	Non-Null Subject Language
NSL	Null Subject Language
OOPT	Oxford Online Placement Test
OPC	Overt Pronoun Constraint
P	Phonological
Pro	Finite null subject
PRO	Nonfinite null subject
SLA	The Field of Second Language Acquisition
stdDev	Standard Deviation
T	Tense
TL	target language
TP	Tense Phrase/Tense Projection
UG	Universal Grammar
UI	upper-intermediate
φ-features	Phi-features
1SG	First singular
1 st	First
2 nd	Second
2SG	Second singular
3 rd	Third

Chapter 1. General Introduction

Ever since the introduction of the Null Subject Parameter by Chomsky in (1982), null subjects have been one of the most intensively investigated phenomena in both theoretical linguistics and language acquisition research. The nature of this linguistic phenomenon, which refers to the cross-linguistic variation in the overt vs. null expression of pronominal subjects, has attracted the attention of many linguists working in the field of second language acquisition (SLA) research. This is because this grammatical variation provides them with an unparalleled opportunity to explore and explain various issues related to the process of second language acquisition (L2A), including access to Universal Grammar (UG), transfer and parameter resetting.

However, research on whether first language (L1) null subject parameter settings transfer in L2A and whether the L1 parameter value can be reset in L2A, particularly by adult learners has produced conflicting and inconclusive answers. These are issues which continue to be debated (e.g., White, 1985; Hilles, 1986; Phinney, 1987; Tsimpli and Roussou, 1991; Al Kasey and Pérez-Leroux, 1998; Licerias and Díaz, 1999; LaFond, 2001 and Sauter, 2002; Judy, 2011; Orfitelli and Grüter, 2014).

The primary goal of this thesis is to contribute to the long-standing debate on the operation of the L1 null subject parameter setting during second language acquisition by providing empirical data from a new comparison of native language speakers acquiring English as a second language (L2). Moreover, in order to get consistent and accurate results that enable us to come to more precise conclusions, special attention is paid to certain methodological problems which include the need to better control certain internal and external factors. I argue that inattention to these factors has resulted in conflicting conclusions about both null subject parameter setting transfer and parameter resetting in L2A.

More specifically, the present empirical study investigates the acquisition of the obligatory overt subject pronouns in L2 English by adult native speakers of Arabic, Finnish and French. These three languages differ from each other in terms of the possibility of allowing null subjects in tensed clauses: in Arabic, null subjects are obligatory unless the pronoun is focused or there is a shifted topic, in Finnish they are optional in some contexts and excluded; in French they are excluded (refer to Chapter 2). Such differences in relation to the possibility of null pronominal subjects in tensed clauses provide suitable comparative grounds to investigate the nature of null subject transfer in adult L2A. This is because, while the Arabic-speaking learners have to acquire a new type of pronouns, namely the weak/unstressed pronouns of English, their French counterparts have such unstressed pronouns in their L1, unlike the Finnish participants who have them but they can be null in certain contexts (see Chapter 2).

There has been no study involving L2 contexts where native speakers of partial null subject languages like Finnish were learning non-null subject languages like English or French; all the previous studies involved only L2 contexts where native speakers of consistent- and/or discourse-null-subject languages like Spanish and Chinese were learning non-null subject languages like English or vice-versa. The present study intends to fill this gap in the literature by presenting empirical data from L1 Finnish-speaking learners of English. Also, no previous study has ever been conducted to examine the acquisition of the Null Subject Parameter in the L2 English of Arabic speakers. Thus, owing to the contrasts between the languages in question, results from this study will provide us with a more accurate picture about whether there is null subject parameter transfer and then resetting in L2A.

The chapters to follow will discuss these points mentioned above in detail. This thesis is organized in six chapters: Chapter 1 (this chapter) illustrates the general goals of this empirical study, namely the gaps in literature it intends to fill. At the beginning of each chapter, a more comprehensive introduction is given to outline the content as well as the structure of that chapter. Chapter 2 introduces and discusses UG. To illustrate how the languages in question vary cross linguistically, the Null Subject Parameter is considered. Then, some of principal approaches to the analysis of null subjects as a theoretical construct are discussed. In order to set the discussion

in the context of language acquisition, this chapter also addresses the process of how children learn their native language with special reference to the phenomenon of early null subjects in child language. Chapter 3 reviews existing literature on the nature of the null subject parameter transfer and resetting in L2 acquisition, particularly by adults. Chapter 4 presents and discusses the study design including detailed information about all aspects of the methodology used in the experiment: the participants, the elicitation techniques and the criteria implemented in the design of these techniques to improve their validity, the test used to assess the participants' levels of proficiency, the data collection procedures and scoring and the coding and data analysis procedures. The specific research questions and hypotheses formulated are also presented in this chapter. The fifth chapter reports and discusses the empirical results in detail. Chapter 6 summarizes the research findings. The study limitations and the suggestions for future research are addressed in this final chapter.

Chapter 2. UG Theory and Language Acquisition: Evidence from the Null Subject Parameter

2.1 Overview of the Chapter

This chapter is meant to introduce the UG-based theory of language acquisition; it focuses on parameters, both as a theoretical construct and in relation to language acquisition. The null subject parameter is used as an example to illustrate how languages vary and to explain how the child's grammar develops and restructures to converge on an adult grammar over the course of time. The process of adult L2A will be discussed in Chapter 3.

The chapter is structured as follows: the second section outlines some of the crucial ideas relevant to language acquisition in generative linguistics such as the notions of competence, performance, critical period, and language faculty. Section 2.3 discusses the nature of the language faculty. This section is divided into four subsections. Subsection 2.3.1 discusses briefly the content of the language faculty from the perspectives of the Principles and Parameters Theory and the Minimalist Program for Linguistic Theory. In doing so, these linguistic theories are introduced. Subsection 2.3.2 describes the contrast between languages with regard to whether they allow empty categories in subject position in finite clauses. Subsection 2.3.3 examines in particular the different contexts where consistent null subject languages and partial null subject languages allow null subjects in comparison with non-null subject languages, with special reference to the four languages of interest in this thesis, namely Arabic, Finnish, French and English. Subsection 2.3.4 considers certain relevant theories that have been put forward to account for this phenomenon. The fourth section is first devoted to discussing how children are hypothesised to acquire their L1; then it examines empirically, in light of findings from the early null subject phenomenon, the content of grammars developed by children at various developmental stages until they acquire the appropriate value for the null subject parameter. The final section highlights the important role of UG to first language acquisition.

2.2 UG and Child Language Acquisition

The field of linguistics has been developing very rapidly. During the last century, several theories have emerged. One which has opened new perspectives in our understanding of both language structure and language acquisition is the theory of Universal Grammar initially proposed by Chomsky in 1957. Chomsky (1975: 29) defines this notion as “a system of principles, conditions, and rules that are elements or properties of all human languages”. The idea is that these principles, conditions, and rules are found in all languages because they are a property of the human mind. With its ultimate aim of integrating “grammar, mind and language at every moment” (Cook and Newson, 2007, p. 11), the primary goal is to “understand the mechanisms which underlie the human ability to build mental grammars” (Hawkins, 2001, p. 1). However, understanding the nature of these internally operating mechanisms “is inseparable from the problem of how it [language] is acquired” (Cook and Newson, 2001, p. 2).¹ Hence, linguists, in order to describe properly such an abstract mental grammatical system need to answer the question: how do children so masterfully acquire the complex knowledge of their native language?

Following the idea which regards language “as a natural phenomenon” (Lenneberg, 1967, p. vii) which should be studied as an “organ of the body” (Chomsky, 2005, p. 133), Chomsky (1957, and much of his subsequent work) proposes what he takes as the most plausible answer to this question, which is that “there is a specific faculty of the mind/brain that is responsible for the use and acquisition of language, a faculty with distinctive characteristics that is apparently unique to the species in essentials and a common endowment of its members, hence a true species property” (Chomsky, 1992, p. 4).² This means that children come to

¹ Although UG guides and constrains child-language acquisition, “it is not, of itself, a theory of acquisition” (White, 1998, p.2). However, “study of what child learners bring to the task of language learning provides insight into the contents of Universal Grammar” (Thomas, 2004, p. 3). This is because the ability to acquire a native language reflects to a great extent some properties of the mind.

² In later work, Hauser, Chomsky and Fitch (2002) divide the language faculty into two subtypes: faculty of language in the broad sense (FLB) and faculty of language in the

the task of language acquisition with prior knowledge as part of genetic endowment that guides them in the course of acquiring their native language. This claim, referred to as the Innateness Hypothesis, is empirically supported in child language acquisition research as pointed out by e.g. Chomsky (1965, 1972, 1981); O'Grady (1997); Lightfoot (1999); Anderson and Lightfoot (2002); Fitch, Hauser and Chomsky (2005) and by the observation that all normal children (1) invariably acquire successfully a remarkably complex grammatical system, and do so (2) at roughly the same pace, (3) following roughly the same developmental process, (4) unconsciously without explicit instruction, (5) despite the fact that the speech input they receive is very often imperfect, in that the speech input they receive is often imperfect, containing false starts, unfinished sentences, and the like, and (6) "do not provide adequate information about complex structures in the language for the child to acquire these on the basis of the input alone" (Lakshmanan, 1994, p. 3). This acquisition phenomenon, where there is a mismatch between the speech input which children are exposed to and their linguistic competence which goes far beyond the impoverished input they receive, is known as the logical problem of language acquisition or the poverty of the stimulus (for further discussion, see Thomas (2002); Sampson (2002); Lasnik and Uriagereka (2002); Scholz and Pullum (2002); Fodor and Crowther (2002) Schwartz and Sprouse 2013).

That a good deal of any native speaker's daily speech is not perfectly grammatical led Chomsky (1965) to distinguish between competence and performance. While competence is "the speaker/hearer's knowledge of his language", performance is "the actual use of language in concrete situation" (Chomsky, 1965, p. 4). Because of the errors caused by performance factors, such

narrow sense (FLN). FLB is an inclusive system which includes all language and communication components some of which are not necessarily unique to humans such as vocalization and communicative behaviour. FLN is a restricted and narrow part of FLB; however, its finite set of elements is unique to humans. Hauser et al. (2002, p. 1571) assume that the "key component of FLN is a computational system (narrow syntax) that generates internal representations and maps them into the sensory-motor interface by the phonological system, and into the conceptual-intentional interface by the (formal) semantic system". See also Fitch, Hauser, and Chomsky (2005).

as slips of the tongue, misinterpretations and processing difficulties due to limited working memory, performance is not on all occasions a perfect reflection of competence.³ This distinction, therefore, is crucial for the theory that “is concerned with what a speaker knows about language as an internal property of human mind rather than something external [the produced utterances]” (Chomsky, 1988, p. 36). So, UG is relevant to competence rather than performance. Chomsky (1986, p. 22) terms this internalised linguistic system (the grammatical competence) as the system of human “I-language”.

Further support for the claim that at least some aspects of language originate in the child’s genetic endowment comes from the idea that “there is a limited developmental period during which it is possible to acquire a language, be it L1 or L2, to normal, native-like levels” (Birdsong, 1999, p. 1).⁴ Studies of child language acquisition lend empirical support to this critical period during brain growth, referred to as the Critical Period Hypothesis (e.g. Lenneberg, 1967; Hurford, 1991; Smith, 2004). The best known example that supports the hypothesis in child language acquisition comes from a child called ‘Genie’ who was totally isolated until age 13. Even after extensive exposure to linguistic input, her subsequent language development was not normal; although she was quite successful in acquiring a large vocabulary, her syntax and morphosyntax never developed beyond a basic level (see Curtiss, 1977). Genie’s syntactic deficits suggest that the critical period for acquiring a native language “holds for the acquisition of grammatical abilities, but not necessarily for all aspects of language” (Fromkin, Rodman and Hyams, 2013, p.

³ When this happens – producing ungrammatical sentences that may violate some principles of UG - this does not mean that the speaker does not know his or her native language. If grammar were not constrained by the principles of UG, then native speakers of any language would be expected to treat grammatical and ungrammatical sentences alike, “since the principle ruling out the ungrammatical sentences would not be available” (White, 2003, p. 29) and they do not. For more extensive discussion, see Grimshaw and Rosen (1990) and White (2003).

⁴ Such critical periods apply to various other living organisms’ innate behaviours that are triggered by specific input (see Bolhuis and Everaert, 2013).

479). Additional evidence for the critical will be discussed in Chapter 3 when the effect of age of first exposure to an L2 is looked at. The following section will discuss the hypothesised nature of the language faculty.

2.3 The Nature of the Language Faculty

2.3.1 Principles and Parameters Theory and the Minimalist Program for Linguistic Theory

Children's built-in language faculty places "limitations on grammars, constraining their form (the inventory of possible grammatical categories, in the broadest sense, i.e., syntactic, semantic, phonological), as well as how they operate (the computational system, principles that the grammar is subject to)" (White, 2008, p. 20). On the other hand there is obviously much variation among the languages of the world as regards the lexicon, phonology, morphology as well as in syntax. The need to resolve the conflict between the conclusion that I-language must be highly constrained and the fact that there is variation across languages gave rise to the Principles and Parameters Theory that was first developed in Chomsky (1981). The central claim of this theory is that the language faculty includes a set of innate universal grammatical principles which define how grammatical operations work. Some of these principles are invariant across languages, while others vary, accounting for the systematic syntactic variation found cross-linguistically. Such variant principles are known as parameters, usually with binary values that were viewed, according to the theory, as predetermined by UG and for which children have to set the value appropriate to the language they are exposed to, based on the linguistic input they encounter.

However, this view of variation has changed since the introduction of the Minimalist approach to UG by Chomsky in (1993), particularly in more recent years: see Chomsky (2005). It has become clear that the notion of a richly specified UG as part of the human genome is unrealistic (Chomsky 2005; Berwick and Chomsky 2011). As a consequence, there is now a more concerted effort to distinguish between universal properties of language that are the result of extragrammatical

factors and those that are the result of UG proper. This approach reduces, therefore, the role of UG, compared to how it was viewed in the eighties and early nineties.⁵

To make the discussion clearer, consider the model of the human language design put forward by Chomsky (2005). This model is referred to as the ‘three factors model’, illustrated as follows:

Factor 1: The genetic endowment (UG), which includes all the universal properties shared by all human languages that need not be learned and cannot be explained by any extralinguistic factors.

Factor 2: Experience (the linguistic data), which leads to variation across languages; the acquirer’s task is to learn, based on the linguistic input he or she receives, which settings are appropriate for the language being acquired for each variant grammatical property.

Factor 3: Principles not specific to the language faculty; these include general properties of computation and general properties of cognition, including learning strategies such as generalising from particular instances to whole categories.

A plausible example of a universal property of human language is the principle which says that every theta role that a predicate can assign must be assigned to one and only one argument (Chomsky 1981). For example, if a predicate can assign an Agent role there must be a determiner phrase (DP) merged with the predicate which can receive this role. Due to principles which are also universal (a universal theta hierarchy; see Baker 1997), this DP will be a subject. However, languages vary with regard to whether or not this subject has to be overtly realised. In particular, languages vary with regard to whether it must be realised in finite/tensed clauses, i.e. whether null subjects are permitted or not. Furthermore, languages not only vary

⁵ For detailed discussion about the Minimalist approach to UG, see Chomsky 1995, 2000, 2001, 2005, 2008, 2013; Sciullo and Boeckx 2011; Boeckx, Horno-Chéliz and Mendivil-Giró 2012.

with regard to whether or not the subject of a tensed clause must be overtly realised, but also languages that permit null subjects vary with regard to the conditions under which null subjects are allowed. In relation to language acquisition, this language-specific grammatical property is learned as a result of Factor 2 (the linguistic input) in the case of child language acquisition; in the case of adult L2A the process of learning this grammatical property can be guided as well by Factor 3 (learning strategies) in addition to Factor 2.

There are still parameters, in the sense that languages vary with respect to a limited number of options. But, unlike classical Principle and Parameter theory, the options are not specified by UG. Instead options arise when UG does not specify a value. Variation occurs because UG is underspecified with respect to various properties (Roberts and Holmberg 2010). The number of options may still be strictly limited, maybe just two, but this is determined by extragrammatical factors. For example, a category may be overt or covert or absent. These are the only logically possible options. If UG requires that the category be present, then the only logically possible options are overt or covert. The language learner has to decide based on primary data which is the option taken in the language being acquired.

The following subsection discusses in detail this type of variation that exists across languages. I will return to the issue of how this variation is acquired by child language learners in section 2.4.

2.3.2 The Null Subject Parameter(s)⁶

One type of grammatical variation among languages can be illustrated by the following contrast between Arabic and Finnish on one hand and English and French on the other hand, as in (1 a, b) and (2 a, b) respectively:

⁶ Holmberg (2010a) points out that “the Null Subject Parameter is often talked about in the singular, even though it is widely recognized that null subjects can be derived in more than one way, and that, therefore, more than one parameter is involved determining

- | | | | |
|-----|----|---|--|
| (1) | a. | Atakallamu ʔal-'arabi:ya
speak-1sg Arabic
‘I speak Arabic’. | [Arabic] |
| | b. | Puhun englantia
speak-1sg English
‘I speak English.’ | [Finnish]

(Holmberg, 2005, p. 539) |
| (2) | a. | *Speak English. | [English] |
| | b. | *Parle français.
speak-1sg French.
‘I speak French.’ | [French]

(Roberts and Holmberg, 2010, p. 4) |

These examples show that certain sentences that are allowed in Arabic and Finnish such as in (1.a, b) are ungrammatical in English and French as in (2.a, b). The verb *atakallamu* in Arabic, for example, can exist without an overt subject, but its English counterpart *speak* requires an explicit referential pronominal or lexical subject; the French finite verb *parle* in (2.b) patterns like its English counterpart *speak* in (2.a) where it obligatorily requires an overt subject pronoun, whereas their Finnish counterpart *puhun* in (1.b) can have null subject just like the Arabic verb *atakallamu* in (1.a).

The initial observation, then, is that languages vary with respect to whether they allow declarative clauses to have null subjects. This linguistic variation across languages is commonly portrayed in terms of a binary condition, with only two possible settings for any given language – it either allows or disallows any finite clause to have a null subject, that is a subject, more specifically a subject, which is syntactically manifested but is not pronounced. This is known as the Null Subject Parameter (Chomsky, 1982). Thus, English and French, among many other languages, are non-null subject languages (henceforth, non-NSL), while Arabic and Finnish, among many other languages, are null subject languages (henceforth, NSLs).

whether subject pronouns can be null or not in a given language.” (Holmberg, 2010a, p. 88).

However, the situation turns out to be much more complex than presented in this simplified scenario. Compare the following Arabic example in (3) with the Finnish one in (4):

(3) Yatakallamu ʔal-'arabi:ya
speak-3sg-m Arabic
'He speaks Arabic.'
(Cook and Newson, 2001: 57)

(4) *(Hän) puhuu englantia
he/she speak-3sg English
'He/she speaks English.'
(Holmberg: 2005: 539)

Although both Arabic and Finnish are considered as NSLs, these examples indicate that the distribution of null subjects varies to some extent. In Arabic the third (3rd) person pronoun as in (3) above must be left unexpressed similar to the first (1st) and second (2nd) person pronouns in a context where there is no emphasis on the subject, neither emphasis by focus nor by topic shift (i.e. when introducing a new topic). However, in Finnish the definite 3rd person pronoun cannot be null in such contexts. More specifically, if the pronoun is null in (4) above, the sentence would have a different interpretation; the null pronoun would be interpreted as a generic pronoun as in (5) below.⁷

(5) Täällä voi puhua englantia.
here can speak English
'One can speak English here.'
(Holmberg, p.c.)

It could be concluded, therefore, that the Null Subject Parameter in its simplest form does not account for all the complex variations that exist among the languages generally. The observation that NSLs vary with regard to conditions under which null subjects are allowed has been observed by comparative linguists for the past 35 years since Rizzi (1982). As a result, NSLs have been classified into several types. For detailed discussion about the typology of NSLs, see among many others Biberauer, Holmberg, Roberts and Sheehan (2010); Holmberg (2005); J. Huang (1984, 1989); Y. Huang (2000) and Rizzi (1982, 1986). The following subsection will

⁷ For further discussion of the null generic subject pronoun in Finnish, see Holmberg (2010b)

be devoted to exclusively examining the different conditions under which null subjects are allowed in two different types of NSLs, known as consistent NSLs and partial NSLs, with special reference to Arabic as an example of the former type and to Finnish as an example of the latter type, in comparison with non-NSLs, with special reference to French and English. However, it must be acknowledged at this stage that despite the significant similarities among the different languages classified as partial NSLs (i.e., Finnish along with Brazilian, Hebrew, Marathi and Portuguese) with regard to the conditions under which null subjects are allowed, there is still some variation among them. Therefore, all the discussion in the following subsection will be relevant to Finnish in particular. For more detailed discussion about the differences in the distribution of null subjects in partial-NSLs, see Biberauer et al. (2010); Holmberg, Nayudu and Sheehan (2009); Shlonsky (2009); Vainikka and Levy (1999).

2.3.3 Similarities and differences among Arabic and

The examples in the previous subsection made it evident that null subjects are allowed in Finnish but under restricted conditions. To be more specific, a 3rd person definite subject pronoun in Finnish must be pronounced unless it is controlled by the closest c-commanding argument. If it is, this pronoun can optionally be null unless it is focused or a shifted topic. However, 1st and 2nd person pronouns can optionally be left unexpressed in basically any context, whether controlled or not, unless the pronoun itself is focused or a shifted topic. This is illustrated in the following examples with an embedded subject pronoun. (6) exemplifies a null 1st person and 2nd person subject. (7) exemplifies the case of an embedded 3rd person subject pronoun controlled by the closest c-commanding argument, the subject of the immediately higher clause.

- (6) (Minä) jään kotiin jos (sinä) pyydät kauniisti
 I remain-1SG home if you ask-2SG nicely
 'I'll stay home if you ask nicely.'

(Based on Vainikka and Levy, 1999. p. 632)

- (7) Pekka₁ väittää että (hän₁) puhuu englantia hyvin.
 Pekka claims that (he) speaks English well
 'Pekka claims that he speaks English well.'

(Holmberg, 2005, p. 539)

I will refer to the condition on null 3rd person subjects as *the control condition*. Consider the following example where the control condition is not satisfied:

- (8) Missä Matti on ollut?
where Matti has been
*On käynyt Pariisissa.
has been in-Paris (Holmberg, 2001: 148)
- (9) Kun hän soitti, *(he) söivät juuri aamiaista
when he called-3SG, (they) ate-3PL just breakfast
'When he called, they were just eating breakfast.'
(Vainikka and Levy, 1999: 636)
- (10) Jari sanoo että lapset uskovat että *(hän) kävi tohtorilla.
Jari says that children believe-3PL that he visited-3SG doctor
'Jari says that the children believe that he went to see a doctor.'
(Holmberg and Sheehan, 2010: 137)

(8) shows that “the anaphoric relation cannot in general extend across independent sentence-boundaries” (Holmberg, 2001, p. 148). This follows from the control condition: the independent clause subject does not have a c-commanding controller. In (9) the embedded plural pronoun *he* ‘they’ must be overt as it cannot be coreferential with, hence cannot be controlled by, the c-commanding matrix subject *hän* ‘he’. The problem with (10) is the fact that the antecedent with the right features is not the closest one; the embedded subject pronoun *hän* is 3rd person singular. The closest c-commanding argument is the plural subject *lapset* ‘the children’, which is not a possible controller.

Typically, the controller of the 3rd person null subject is in the next higher clause. Exceptions to this can occur only when the next clause up contains no argument. In this case, the 3rd person null subject can be coreferential with an argument that is more than one clause up. Such a possibility is illustrated by the following example:

- (11) Marja₁ sanoo että on varmaa että (hän₁) saa ensi vuonna ylennyksen.
Marja says that is certain that she gets next year promotion
'Marja says that it's certain that she will get a promotion next year.'
(Holmberg and Sheehan, 2010, p. 20)

In contrast to Finnish which does not require obligatory null subjects in finite clauses but allows them optionally where possible, Arabic requires all of the three person pronouns in singular, dual and plural to be left unexpressed in any finite

context unless the pronoun is focused or a shifted topic. Consider the following examples:

- (12) a. sa-ađhabu haythumaa *(anta) tađhab.
FUT-go-1SG.NOM wherever (you) go-2SG.NOM
 'I will go wherever you go. (Based on Alqurashi, 2010: 14)
- b. 'aqsama Zayd-un an sa-yaghlibu *(huwa) Amr-an
swore-3SG.M. Zayd-NOM that FUT-beat-3SG.M (he) Amr-ACC.
 'Zayd swore that he will beat Amir.' (Based on Al-Seghayar, 1997 :3)
- c. sa-ađhabu řindamaa *(huwa) yařawadu
FUT-go-1SG.NOM when (he) PRES-return-3SG.M
 I will go when he returns.

All these sentences in (12) are ill-formed unless the subject pronoun is focused or is a shifted topic. Note one exception to the condition that a subject pronoun which is not focused or a shifted topic must be null in Standard Arabic or in the Saudi Arabic dialect (the variety relevant to this thesis) occurs when the subject following *řinna* and *řanna* 'that' is a pronoun; in such a case the subject is realized as a clitic on the complementiser and is assigned accusative (see Fassi Fehri, 1993, p. 98; Aoun, Benmamoun and Choueiri, 2010, p. 14; Johns, 2007, p. 129). This is illustrated by (13.a):

- (13) a. qà-a řinna-hu dżà?-a
said-3SG.NOM that-3SG.M.ACC came-3SG.M
 'He said that he came.' (Johns, 2007: 129)

Another exception to this condition occurs in verbless sentences. In such constructions the subject pronouns must be overtly expressed (see Fassi Fehri, 2009, 2011; Aoun et al, 2010; Hole, 2004). Consider the following example: Without the overt subject, the sentence is ungrammatical.

- (13) b. *(?ana) mudarris
 I teacher
 'I am a teacher' (Holes, 2004: 183)

The discussion in this subsection and the previous one is meant to make it clear that null subjects in Finnish are optional in some contexts and excluded in other contexts, whereas they are obligatory in Arabic and excluded in English and French. These descriptive observations can be summarised as follows:

- a) A subject pronoun in Finnish tensed clauses can optionally be null unless:
 i) it is focused or a shifted topic, or

- ii) it is 3rd person and is not locally controlled.
- b) A subject pronoun in Arabic tensed clauses must be null unless:
 - i) it is focused or a shifted topic, or
 - ii) it is subjacent to an λ -type complementiser, or
 - iii) it is a subject of a verbless clause.
- c) A subject pronoun in English or French tensed clauses must be overt.

It should be mentioned at this point before completing this subsection that despite the fact that both English and French require subject pronouns to be overtly expressed, they do allow null subjects in a restricted context, namely, in matrix clauses, strictly in sentence-initial position, and under certain discourse and register conditions. This is particularly common with 1st person singular subjects. However, “embedded subjects consistently remain overt” (Haegeman, 2000, p. 138). Consider the following examples:

- (14) a. Can't find my pen.
 b. Think I left it at home.
 c. *Think left it at home. (Radford, 2004, p. 107)
- (15) a. Elle est alsacienne. Parait intelligente.
 She is Alsatian. Seems intelligent.
 (Leautaud 1988: 48, from Roberts and Holmberg, 2010: 5)

Even in sentence-initial position in root contexts, including when the subject is 1st person singular, a null pronoun is not always licit. This is illustrated by the following examples from English where the root subject has to be overt: Consider the following examples:

- (16) a. He is tired. b. *Is tired. (Radford, 2004: 107)
 c. I am tired. d. *Am tired.

Such distributional constraint on null subjects in English and French suggests that they are “derived by a mechanism different from the one which derives null subjects in consistent and partial NSLs” (Holmberg: 2010a, p. 90). It has been argued that such null subjects allowed in main clauses in non-NSLs are derived by topic drop; for more detailed discussion see among many others, Haegeman (1990, 1997, 2000), Rizzi (1992), Cote (1996), Haegeman and Ihsane (2001), Rodrigues (2002), Radford (2004) and Roberts and Holmberg (2010).

The next subsection seeks to explain the cause(s) of this observed fact: why some language-specific grammars permit null subject while others do not. It presents some of the different theories that have been proposed to account for the existence of the null subject phenomenon.

2.3.4 The syntax of the null-subject phenomenon: licensing and identification within the agreement-based analysis

A number of theoretical questions can be raised on the basis of the previous discussion in relation to the syntax of null subject phenomena, such as:

1. *Why do some languages allow a null subject in tensed sentences and others do not? How can its content be identified or recovered?*
2. *Why do null subjects occur more freely in consistent NSLs compared to partial NSLs? What is the reason for the split between 1st and 2nd person on the one hand, and the 3rd person on the other hand, in partial NSLs?*

Several proposals have been offered to answer these questions. However, I will not undertake a critical review of all the different theories, as they have evolved over the past 30-40 years. However, the exact nature of this syntactic variation is still subject to heavy debate among linguists and has been being revised ever since Perlmutter (1971). This section will, therefore, explore only some proposals that can explain the syntactic mechanisms by which a null subject is licensed in the NSLs covered in the thesis.⁸ Namely, it will focus in particular on certain agreement-related proposals that are assumed to play a role in deriving null subjects in Arabic and Finnish.

It has been observed since the works of Perlmutter (1971) that languages with rich subject-verb agreement systems tend to allow finite clauses not to have a pronominal subject. This observation has led a number of linguists (e.g. Chomsky 1981, 1982; Jaeggli 1982; Rizzi 1982; Jaeggli and Safir 1989; Hyams 1986, among many others) to argue that the possibility of having referential null subjects (referred

⁸ The interested reader is referred to Biberauer et al. (2010), which is a relatively up-to-date reference book that reviews all existing accounts.

to as ‘*pro*’ in Chomsky 1982) in languages correlates with the range of verbal agreement morphology they realise. Chomsky (1982, p. 241) argued that "the intuitive idea is that where there is an overt agreement, the subject can be dropped, since the deletion is recoverable" by the Phi-features in T. Phi-features (ϕ -features) include the person and number features (and grammatical gender features in some languages).

However, such an agreement-based account, though it covers a great deal of empirical ground by explaining the cross-linguistic differences between the canonical null subject languages (Arabic as well as Italian and Spanish) and the canonical non-null subject languages (English and French), runs into difficulty when extending it to other types of null subject languages, i.e., discourse pro-drop languages (Chinese), Partial NSLs (Finnish) or other type of languages which allow expletive null subjects, but not referential ones (German).⁹

For example, in formal Finnish, T includes the ϕ -features required to identify the null subject according to this agreement-based account, yet they cannot do so as freely as in Arabic.¹⁰ This is illustrated in Table 2.1 where the verb *puhu* ‘speak’ in Finnish inflects for person and number features similar to its Arabic counterpart:

Table 2-1. Subject-verb agreement morphology of the present tense of the verb *puhu* ‘speak’ and its Arabic counterpart

	Finnish	Arabic
1 st person singular	Puhun	ʔa-takallam(u)
2 nd person singular	Puhut	Ta-takallam(u) (M) Ta-takallam-iin(a) (F)

⁹ For the various proposals accounting for null subjects in the different types of NSLs and their assumed associated clusters of syntactic and morphological properties, see Chomsky (1982); Rizzi (1982, 1986); Jaeggli and Safir (1989); J. Huang (1984); Y. Huang (2000); Holmberg (2005, 2010a); Biberauer et al. (2010) among many others.

¹⁰ It should be noted that most varieties of colloquial Finnish, unlike formal Finnish, do not make a morphological distinction between 3rd person singular and plural. They both have the form of the singular in Table 2.1. It is likely that the loss of the number distinction in colloquial Finnish is, at least in part, a consequence of the absence of pro drop with 3rd person pronouns.

3 rd person singular	Puhuu	Ya-takallam(u) (M) Ta-takallam(u) (F)
1 st person plural	Puhumme	Na-takallam(u)
2 nd person plural	Puhutte	Ta-takallam-uun(a) (M) Ta-takallam-na (F)
3 rd person plural	Puhuvat	Ya-takallam-uun(a) (M) Ya-takallam-na (F)

This raises the obvious question of why referential null subjects are equally possible in all three persons in Arabic but is restricted to the first two persons in Finnish. Holmberg (2005, 2010) argues that null subjects are derived in two ways:

- a) In languages with subject-verb agreement, that is languages which have unvalued phi-features (ϕ -features) in T, T will enter an agreement relation (an Agree-relation, in Chomsky's 2000, 2001 terms) with the subject DP. Through this relation the subject will get nominative case and T will have its ϕ -features valued. If T has a rich enough set of ϕ -features, so that all the features of a pronominal subject are represented in T, then the subject can be, and even must be, deleted, i.e. must be null, being formally a copy of T. If T's ϕ -feature set contains a D-feature (D=definite) as in Arabic, a definite subject pronoun will be null, by this mechanism. If T does not have a D-feature (as in Finnish), a definite subject pronoun cannot be deleted by this mechanism, only an indefinite one, interpreted as generic.
- b) A subject DP in the specifier of tense phrase (TP) position may be null just by virtue of having an antecedent in the linguistic context which provides it with a referential index. There is variation among languages regarding the contextual conditions. In Arabic, the antecedent can be a DP in a higher clause, or it can be a DP in a preceding independent sentence. In Finnish, the contextual conditions are different for 1st and 2nd person subjects on the one hand, and 3rd person subjects on the other. A 1st or 2nd person subject can be null without any linguistic antecedent. A 3rd person subject can, and must be, null if it is generic, but if it is to have a definite reading, it must either be spelled out or have an antecedent in the next higher clause: it must be controlled, in the sense of Holmberg (2005, 2010a), Holmberg, Nayudu and Sheehan (2009).

Intuitively, a difference between 1st /2nd person pronouns and 3rd person pronouns is that 1st and 2nd person pronouns always 'have an antecedent' in the

discourse, namely the speaker and the addressee. Holmberg (2010a) proposes formalizing this as in Sigurðsson (2004), where the speaker and the hearer are syntactically represented as features in the C-domain of every main clause. In these terms, a 1st and 2nd person subject will always have a linguistic antecedent. A 3rd person subject will have a linguistic antecedent only if there is a c-commanding DP in the next higher clause. As a consequence, Finnish allows 3rd person definite null subjects only in embedded clauses.

Mechanism (b) is independent of agreement. Why do not all languages avail themselves of this possibility? There is a requirement in some languages that the specifier of TP position, the grammatical subject position, needs to be spelled out/pronounced, regardless of linguistic context and regardless of the ϕ -feature content of T. (Holmberg 2010a, p. 114-116). English and French are such languages. Holmberg (2010a, p. 114-116) formalises this as a parameter:

(17) T has or doesn't have a P-feature (P=phonological EPP).

The effect of the feature P is forcing the specifier of T to have spelled out content, typically a subject DP moved there, but in the absence of a thematic subject it can be an expletive. In some languages it can be some other constituent of the sentence moved there (see Holmberg, 2010a).

According to this parameter in (17) above, it can be stated that T in Arabic and Finnish does not have a P-feature in T, whereas T in English and French does. So, this is the parameter that needs to be reset by Arabic and Finnish learners of English in the present study.

In the following I will refer to this parameter as [\pm null subject]. Languages with P in T, in Holmberg's (2010a) terms, have the value [-null subject], languages without P in T have the value [+null subject].

2.4 Null subjects in first language acquisition

Having theorised that all normal children are born with an in-built language faculty with a set of finite universal principles, it is assumed that these principles constitute the starting point of language acquisition (G_0) for all children. In other words, these

principles form “the initial state of the language learner, hence the basis on which knowledge of language develops” (Chomsky 1988, p. 69). Therefore, it is expected that children at the G_0 stage start to learn the variant grammatical properties in a largely uniform way.

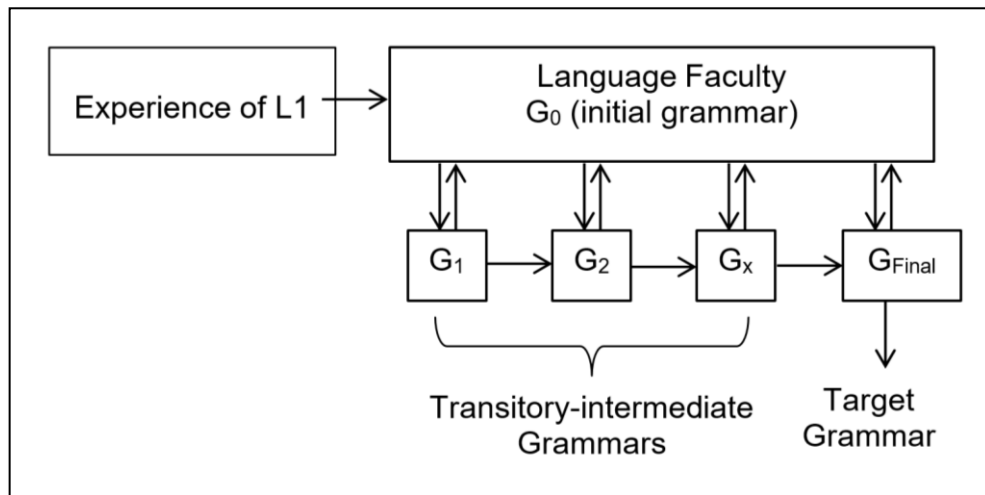


Figure 2-1. Model of child L1 acquisition

To illustrate this uniformity notion in children’s cross-linguistic language development, let’s look at the early null subjects in L1 acquisition. It has been widely observed that children, regardless of whether their target language is a NSL or not, pass through certain transitory stages in their grammatical development where they initially produce finite sentences with no overt subjects (see R. Brown, 1973; L. Bloom, Lightbown and Hood, 1975; Valian, 1991; Wang, Lillo-Martin, Best and Levitt, 1992; Pierce, 1994; Rizzi, 1994; Rasetti, 2000; Valian and Eisenberg, 1996; Hamann and Plunkett 1998; Hamann, Rizzi and Frauenfelder 1996, and the references cited below). The following examples in (18) illustrate these early null subjects; note that the relevant adult languages fall into different groups in terms of licensing null subjects – English, French and Danish are non-NSLs, Italian and Japanese are NSLs and Hebrew is a partial NSL.

(18)

- | | |
|---|--|
| <p>a. English
Want more apples
*(I) want more apples.
(Hyams 1986, after L. Bloom, 1970)</p> | <p>b. French
Oter tout ta.
empty all that
*(I) empty all that.
(Hamann and Plunkett, 1998)</p> |
| <p>c. Danish
Ikke køre traktor.
Not drive tractor.
*(I, you, he) doesn’t drive the tractor.
(Hamann and Plunkett, 1998)</p> | <p>d. Italian
Butta via.
(he) throws away
(Serratrice, 2005)</p> |

e. Hebrew
 Shata.
 drank/3msg DirObj this
 *(He) drank this.
 (Levy and Vainikka, 1999)

f. Japanese
 Hamu taberu.
 ham eat
 'I'll eat ham.
 (Hirakawa, 1993)

However, despite this apparent similarity in the initially developed grammar, it has been noticed that the actual percentage of subject drop produced by children, as they pass from the initial state (G_0) through the multiple transitory mental grammatical states (G_1 , G_2 , G_x), varies considerably based on a number of factors: the target language (the input), age of the acquirer and the produced syntactic construction (Bates, 1976; Valian, 1991, and Aronoff, 2003). Table 2.2 below illustrates the large differences in rate of null subjects across child languages where the age ranges of the children are similar.

Table 2-2. Percentages of Null Subjects across Child Languages

Child L1	Age	Subject drop rates	Source
English	2;03 - 2;08	15%	Hyams and Wexler (1993: 426)
French	2;03 - 2;07	38%	Jakubowicz, Miller, Riemer and Rigaut (1997, p. 335)
Japanese	2;03 - 2;06	79%	Hirakawa (1993:43)
Chinese	2;00 - 2;05	56%	Wang et al. (1992:238)
Spanish	2;5	66.3%	Bel (2003: 9)

Note the children acquiring null subject languages (Japanese, Chinese, Spanish) produce finite sentences with no overt subjects with much greater frequency compared to their counterparts acquiring non-null subject languages (English, French) during approximately the same age. Such differences could be attributed to the different properties of the input the children receive; the former group of languages are known to make massive use of null pronouns, while the latter prohibit or highly restrict null subjects.

It should be mentioned at this point that it has been observed that English-speaking children and Inuktitut-speaking children at this stage omit 1st and 2nd person pronouns more frequently than the 3rd person (see Valian and Eisenberg 1990,

Hughes and Allen, 2006).¹¹ Prévost (2009, p. 135) finds in case of French children that the first pronouns to “emerge are third person singular pronouns, such as *il*, *elle*”. However, over the course of time, and only on the basis of positive evidence, this divergent transitory intermediate grammar starts to converge on the target-adult grammar. Approximately at age three, children arrive at the appropriate value for the given grammatical property; that is, children learning a language like English acquire the [-null-subject], children learning a language like Arabic acquire [+null-subject] in order to construct the target core grammar.

An important question to be raised on the basis of the above discussion is: what is the nature of these early null subjects? The phenomenon of child null subjects has been accounted for under two approaches: a competence-based approach and a performance-based approach. Within each approach, different accounts have been offered (see, among others, P. Bloom, 1990; Hyams, 1986, 1992; Rizzi, 1994, 2000; Valian, 1990, 1991; Valian and Eisenberg, 1996; Bromberg and Wexler, 1995; Orfitelli and Hyams 2008). However, the exact nature of these early null subjects is still subject to debate among linguists.¹²

Hyams (1986) originally proposed that children’s early null subjects are *pro*, licensed by the same mechanism that licenses null subjects in Arabic-type-languages. She argued in case of children acquiring non-NSLs, such as English, that once they discover the impoverished agreement system, null subjects are blocked. Many researchers, including Rizzi (1998) and Roeper and Rohrbacher (2000), argued following Hyams (1986) that children set the value of the null subject parameter once they acquire the agreement morphology. In a later article and in order to deal with the flaws of her pro-drop model,¹³ Hyams (1992) assumes that child null

¹¹ Inuktitut is primarily spoken in native populations in Canada.

¹² The interested reader is referred to Guasti (2002) and to Hyams (2011) for an historical review.

¹³ For reasons of space, these problems will not be presented here; however, for detailed criticism, see Haegeman, 1990; Rizzi, 1994; Valian, 1990).

subjects are topic-drop, licensed via T and identified by a discourse topic chain. In this way, the early null subjects in child grammar of English (or Arabic, etc.) resemble that of null subjects in topic-drop languages like Chinese (see J. Huang, 1989). If this is the case, then it would be expected that null subjects in child English have the same distribution as in adult Chinese. However, Hughes and Allen (2008) observed that children omit null subjects even in cases when the referent cannot be identified from the discourse. Moreover, Rizzi (1994, p. 155) observed that the children's null subjects are structurally "limited to the initial position, the specifier of the root" – a phenomenon similar to null subjects found in the diary registers of adult English speech (see subsection 2.3.3). Therefore, Rizzi (1994, 2002, 2005a,b) proposed a truncation analysis for such an empty antecedentless category, assuming that early subject drop is root subject drop in that children during the null subject phase, for syntactic developmental reasons, produce incomplete tree structures where the specifier of the root is not merged.

The alternative approach to accounting for the early null subjects is that there are performance-deficit explanations for the phenomenon. For example, P. Bloom (1990) and Aronoff (2003) argue that the child's grammar is similar to that of the adults; however, for processing difficulties, caused by the child's limited working memory or syntactic complexity, omissions occur in the child's production. Accordingly, Bloom claims that children tend to omit subjects from longer and complex utterances more often than from shorter ones; this is referred to as the VP length analysis to early null subjects (see also Valian, 1990, 1991 for a similar claim). Similarly, Allen and Schroeder (2003), Clancy (1993), Gürkanli, Nakupoglu and Özyürek (2007) and Guerriero, Cooper, Oshima-Takane and Kuriyama (2001) found that children show a higher null subject rate when verbs are transitive compared to intransitive; this is because transitive verbs are associated with given information and therefore are in longer sentences. This account is supported empirically by data showing that children drop other constituents in their speech, in addition to null subjects, such as auxiliaries and modals (see, P. Bloom, 1990). However, a counterargument is that children frequently also omit subjects in very simple utterances such as: *want daddy* (see Radford, 1986, 1990). According to the performance-limitation explanation, children are predicted to overtly spell out the subjects in such simple clauses, yet, Rizzi (2002) specifically argues that this early null subject phenomenon

is not attested in embedded clauses, simply because such complex structures emerge later at a stage (at age three) when children have arrived at the appropriate setting for this grammatical property (see the discussion in subsection: 4.3.2.1.2).

2.5 The role of generative grammar in the study of first language acquisition: concluding remarks

The UG-based theory of language acquisition provides an account of how the child's first language development proceeds. It provides linguists with a way to understand the question of how children acquire the grammar of their native language in such a rapid and uniform fashion, based on the impoverished input they are exposed to.

The view that the child's grammatical learning is constrained by invariant principles simplifies the task of acquisition by reducing the syntactic learning required from the child. Since the innately endowed principles do not have to be learned, the child's only grammatical learning is to arrive at the appropriate value for each variant grammatical property in the language being acquired.¹⁴ These conclusions still stand, even in the light of recent developments in generative linguistics. The role of UG may be reduced, in favour of extragrammatical factors, but acquisition of syntax is still largely a matter of choosing between options provided by an underspecified UG.

The next chapter will explore the implications of the theory for the study of second language acquisition (SLA).

¹⁴ Language acquisition also involves lexical learning. Chomsky (1995, P. 28) defines the process of language acquisition as "the acquisition of lexical items, fixing of parameters, and perhaps maturation of principles".

Chapter 3. The Role of UG and the Status of L1 Null-Subject Parameter-Setting Transfer in L2 Acquisition

3.1 Introduction

Having introduced in Chapter 2 the role of generative linguistics in first-language acquisition research, this chapter starts with how UG-based SLA research emerged and its major developments. To show the theoretical and practical significance of UG-based SLA research, it is helpful to retrace the recent history of SLA research. This will be the focus of section 3.2, below; it is important to mention at this point that this section is not intended to give a comprehensive, chronological introduction to the development of the field (see Selinker, 1992; Braidì, 1999; Hawkins, 2001; White, 2003; Thomas, 2004). Instead, it is intended chiefly to address the concepts, observed facts, and controversial ideas about L2A that are relevant to the topic of this thesis: null-subject parameter transfer in L2A.

3.2 Universal Grammar and Second-Language Acquisition

The relationship between generative linguistics and SLA can be said to have started in 1967 with the publication of (*The Significance of Learners' Errors*), in which Corder rejected the contrastive analysis hypothesis (CAH) of Lado (1957) that treats all learners' errors as the result of carrying over bad L1 habits into the L2. Corder instead wondered whether L2 learners' errors are systematic, whether they are actually rule-governed behaviour. In other words, he questioned whether the processes of first- and second-language acquisition are essentially the same, guided by the same language-acquisition mechanism. Corder's insight about the importance of errors and their analysis served to liberate SLA research from the earlier ties to structural linguistics, behavioural psychology, and contrastive analysis (for more detailed discussion, refer to Thomas, 2004). However, the actual birth of so-called generative SLA research began five years after Corder's article with the publication of another seminal article, "Interlanguage". In this article (1972), Selinker shifted the focus of

SLA research to the learner's development of grammar as a whole rather than focusing only on his or her non-target-like performance errors at a particular time. He argued that the L2 learner constructs a UG-constrained grammatical system that may be different from the syntax of his or her first and second languages – a non-defective developing system with its own rules. He refers to this L2 linguistic system as *interlanguage* (IL).

Selinker's profound idea has been significant and remains so, as the ultimate goal of SLA research has been the same ever since: to describe the nature of the IL by explaining how an L2 is acquired and why it is acquired in that way. Yet because linguistic theory provides the baseline for SLA research and has been under regular theoretical revisions ever since, the questions asked about the nature of the IL are regularly changing. These changes can offer new perspectives on language acquisition data and thus allow researchers to draw deeper conclusions. While these conclusions might not be definite, the application of generative theory to SLA studies among researchers sparks renewed interest about the problem of L2A. For example, much of the research in SLA done in the 1970s attempted to provide evidence that SLA is inherently systematic and independent of L1. To address these issues, many L2 researchers at that time mirrored the research that had been conducted in L1 acquisition, basically to determine whether SLA is similar to or different from L1 acquisition. Dulay and Burt (1973, 1974), for instance, conducted morpheme studies similar to that of Brown (1973) in L1 acquisition, who found evidence that all child L1 learners acquire grammatical morphemes in a remarkably similar fixed order. Dulay and Burt found the same for child L2 learners and discovered that this order was somewhat distinct from the L1 acquisitional order reported in Brown's study. Based on such studies, Dulay, Burt, and Krashen (1982, p. 207–209) concluded that "children of different language backgrounds learning English in a variety of host country environments acquire eleven grammatical morphemes in a similar order". Similar acquisitional sequences were noticed regardless of the learning context, whether in a classroom or in a natural acquisitional setting (see Lightbown, 1987; Makino, 1993). Similar findings amongst adult L2 learners of English from different L1 backgrounds were reported by Bailey, Madden, and Krashen (1974). Accordingly, they argued that "adults process linguistic data in ways similar to younger learners" (Bailey et al., 1974, p. 240).

This systematicity in early L2 acquisition studies was seen as evidence supporting the view that L2 learners are innately guided by the same UG principles that guide L1 learners. Nevertheless, the findings that show that the emergence order for the L2 grammatical morphemes differs from that found in L1 acquisition suggested that these principles could be different to some extent (i.e., Dulay and Burt, 1973, 1974; Brown, 1973; Bailey et al., 1974). This suggestion, with the development of the principles and parameters theory, shifted the interest of researchers during the 1980s and early 1990s from whether UG continues to be available in SLA to what kind of UG access is still available for L2 learners. The principles and parameters theory enables researchers to apply “hypotheses about principles and parameters of UG to observable patterns of L2 development . . . to confirm or disconfirm their involvement” (Hawkins, 2001, p. 10), and there is still disagreement among researchers. Basically, three conflicting proposals have been suggested concerning adult L2 accessibility to UG: (a) no access to UG, (b) full access to UG, and (c) partial access to UG.¹⁵ These proposals concern the initial state representations those L2 learners start out with before they receive any L2 input. They can be summarised briefly as follows:¹⁶

The no-access hypothesis assumes that L2 learners no longer have access to UG after passing the critical period, following the critical period hypothesis (CPH),

¹⁵ Note that different terminologies are used in the literature to refer to these positions. For instance, terms such as direct access and indirect access were used to refer to full and partial access, respectively, and terms such as global impairment and local impairment were used to refer to the no-access position and the partial-access position, respectively.

¹⁶ For more detailed discussion, refer to White (2003).

which assumes that the individual's biological maturation affects the language faculty (see section 2.2).¹⁷ Proponents of this view (e.g., Clahsen and Muysken, 1986; Clahsen, 1988; Bley-Vroman, 1989; Schachter, 1990; Meisel, 1991, 1997; Clahsen and Hong, 1995) claim that the process of adult L2A, unlike child language acquisition, is guided by aspects of the mind other than UG – namely, general learning strategies such as problem-solving abilities, instruction, and trial and error. Advocates of this view argue that adult SLA differs fundamentally from child L1 acquisition in terms of the time required to learn the target language, the path of acquisition (see Clahsen and Muysken, 1986), and variability of success. Unlike L1 acquisition, adult L2A is usually rather slow and not uniform, and the final level reached even by successful learners very often stops short of native-like proficiency. Further evidence used in support of this view that IL is not constrained by UG comes from the documented findings that the syntactic and morphological properties associated with a single parameter do not form a parametric cluster in L2A but are acquired individually, rather than simultaneously (see Clahsen and Hong, 1995). Note that the absence of clustering acquisition in IL (see research related to L2A of the null-subject parameter) does not mean that ILs are not constrained by UG (Liceras, 1989; White, 1989, 2003). Ayoun (2000) points out that “parameter resetting may be evidenced by a partial clustering of properties as a result of progressive manifestation of parameter-setting properties” (p. 79).

The extreme opponents of this no-access position (e.g., Dulay et al., 1982; Flynn, 1987, 1996; Epstein, Flynn, and Martohardjono, 1996) argued exactly the opposite; they explained that UG is fully available for L2 learners in the initial state and at every point of acquisition, just as it is for L1 acquirers. This view is often

¹⁷ There is disagreement among researchers (i.e., Johnson and Newport, 1989; Bialystok and Miller, 1999; deKeyser, 2000; Birdsong and Molis, 2001; Birdsong, 2004) about the exact time when child language acquisition – whether a first, a second, a third, or any other language – ends and adult SLA starts. However, some researchers (e.g., Hawkins, 2001a) claim that the boundary between them can be found somewhere between the ages of seven and nine. For a detailed discussion about the effects of age of first L2 exposure, refer to Patkowski (1980), Hyltenstam and Abrahamsson (2003), and Birdsong (2004).

referred to in the literature as the *full-access hypothesis*. It is argued under this hypothesis that an L2 is acquired only on the basis of interaction between UG and L2 input. The main evidence used to support this assumption comes from studies of the logical problem of L2 acquisition – the phenomenon in which L2 learners master abstract properties that neither are instantiated in the L1 nor can be induced from the input they receive. As an illustration, consider the Overt Pronoun Constraint (Montalbetti 1984) in the context of L2A research. For example, Perez-Leroux and Glass (1997) and Kanno (1998) found that native speakers of English who were L2 learners of non-NSLs (Spanish and Japanese) followed the Overt Pronoun Constraint (OPC), disallowing quantified antecedents only for overt subject pronouns but not for null pronouns. Perez-Leroux and Glass (1997) and Kanno (1998) argue that such distinctions in their performances could not have been derived from the L2 input or the L1 grammar alone. Note while it is possible in English for embedded subject pronouns to have quantified or *wh*-antecedents in the main clauses, in Spanish and Japanese only null pronouns can do so, and the overt embedded subject pronouns cannot receive bound variable interpretation (For more detailed discussion, see White 2003).

In contrast to the no-access and full-access hypotheses, the partial-access hypothesis recognises the involvement of the L1 grammar in defining the L2 initial-state grammar. Falling under this hypothesis are two scenarios that share the assumption that UG is accessible via prior grammatical knowledge, which forms the learner's initial representation of the L2. This initial representation is used to analyse the L2 input the learner receives and therefore uses in the construction of the target grammar. However, the two scenarios differ in the extent to which they posit that UG innate linguistic knowledge can be accessible. One scenario is represented by Tsimpli and Roussou (1991), Vainikka and Young-Scholten (1994, 1996a, 1996b, 2011), Eubank (1996), Eubank and Grace (1996), and Beck (1998), who argue that the role of L1 transfer is restricted in L2A. Tsimpli and Roussou (1991) argue that only the parameter values that have been activated in the L1 are available to the L2 learner. Vainikka and Young-Scholten (1994, 1996a), on the other hand, argue that only lexical categories transfer and that the initial state lacks functional categories altogether. This view, originally known as the Minimal Trees Hypothesis and, including Structure Building, has subsequently been referred to in Vainikka and Young-Scholten (2011) as Organic Grammar, was developed based on the Weak Continuity

Hypothesis for L1 acquisition, which assumes that child's initial grammar lacks functional categories, instead containing only lexical categories (see Clahsen, Penke, and Parodi, 1993/1994; Clahsen, Eisenbeiss, and Vainikka, 1994; Clahsen, Eisenbeiss, and Penke, 1996). Other researchers (i.e., White, 1985; du Plessis, Solin, Travis, and White, 1987; Cook, 1994) have, on the contrary, proposed that all parameters of UG are indirectly accessible via learners' L1 grammar. They argue that when the initially L1 transferred grammar is insufficient for the learning task, UG is fully accessible.

All of these L2 UG-accessibility hypotheses, however, have empirically shown to be problematic, though not equally problematic (see White, 2003). To address some of these challenges with regard to each hypothesis, it is better to consider some of predictions of each about L2 end-state IL based on its theoretical relationship with the initial state hypothesised (for a detailed discussion, see White, 2000). White (2008) argues that "initial-state theories necessarily have implications for the nature of representation during the course of development, as well as for end-state representation" (p. 33). This end-state IL grammar is also referred to in the literature as *ultimate attainment*, *the steady-state*, or *the final-state*.

It can be predicted—for example, under the no-access hypothesis of the initial-state since UG is assumed not to be available in L2A—that the L2 final-state grammar is not only incomplete and divergent in relation to the target grammar but it is also sometimes "wild", in the sense that it does not always obey the general generative rules of natural human languages. This implies that native-likeness in L2A cannot be attained at all (Coppieters, 1987; Schachter, 1989, 1990; Bley-Vroman, 1989; Clahsen and Hong, 1995; Neeleman and Weerman, 1997). However, it has been observed that "language mastery is not often the outcome of SLA" (Larsen-Freeman and Long, 1991, p. 153) and "very few L2 learners appear to be fully successful in the way that native speakers are" (Towell and Hawkins, 1994, p. 14).¹⁸ This obviously raises the question of how it is possible for those talented few L2 learners to exhibit native-like

¹⁸ For a review of the literature on near-native-likeness, see Birdsong (1992), White and Genesee (1996), and Sorace (2003).

attainment if their ILs are not UG-constrained and were not restricted to properties found in natural languages. Mitchell and Myles (2004) argue that “UG is a theory of natural languages; claiming it plays no part in SLA would mean claiming that second languages are not natural grammars” (p. 77). Such a belief has led the full-UG-access-based theorists, since they assume ILs are the product of the same cognitive mechanism that is responsible for first-language grammar, to argue that, like the initial-state, steady-state ILs cannot be wild (e.g., White, 2003; Mitchell and Myles, 2004).¹⁹ Native-like competence at the final state is expected; “there should be no evidence of incompleteness, divergence, indeterminacy, non-native optionality” (Birdsong, 2004, p. 94). In other words, L2 end-state grammars should correspond exactly to the target grammar if there is sufficient input.

However, various empirical studies have provided evidence against the view predicting that L2 end-state grammars should correspond exactly to the target grammar. It has been widely observed that the overwhelming majority of L2 learners’ end-state ILs fossilise (cease to develop) at some particular developmental stage that is not native-like (refer to, among many others, Selinker, 1992; Long, 1997, 2003; Han, 2004). This permanent intermediate IL state is the predicted end-state grammatical system assumed by the partial-UG-access theorists (i.e., Tsimpli and Roussou, 1991), who believe that only the grammatical features that have been activated in the L1 are available to the L2 learner. Namely, this hypothesis predicts that the end-state grammar, even though it is subject to UG constraints, is necessarily different from the grammar of the L2 since resetting parameter values is claimed to be impossible in accordance with the no-parameter-resetting hypothesis (i.e., Tsimpli and Roussou, 1991; Smith and Tsimpli, 1995; Hawkins and Chan, 1997), which posits that parameter values not realised in the L2 learner’s L1 are not accessible and that L1 parameters values therefore cannot be reset. Other researchers (e.g, Pollock, 1989; White, 1992; Al-Kasey and Pérez-Leroux, 1998), however, share the belief that the L1 final state is the L2 initial state. They argue that L2 learners converge at the final state on the target language grammar, and they show evidence that L2 learners are able to reset the L1

¹⁹ This viewpoint that ILs cannot be wild is held by the other UG access hypotheses – the partial-access hypotheses.

parameter values to those appropriate for L2 with time and increased proficiency in L2A even if these are values not recognised in their L1. This supports the parameter resetting hypothesis of Schwartz and Sprouse (1994, 1996) and Epstein et al. (1996). These researchers argue that the fact that learners struggle with L2 learning does not mean that learners will be unsuccessful in resetting parameter values. Organic Grammar (the Minimal Trees hypothesis) shares this view about the end-state IL grammar and assumes that only part of L1 grammar constitutes the initial state (L1 functional categories are not present); it predicts that functional categories emerge gradually in response to the L2 input until they are fully mastered. In other words, this hypothesis also predicts that L2 end-state grammars should correspond exactly to the target grammar if there is sufficient input. It is clear so far that all these hypotheses about the initial state (and end state) of L2A have empirically shown to be problematic.

An alternative widely accepted hypothesis among researchers is the full-transfer/full-access hypothesis (FT/FA) of Schwartz and Sprouse (1994, 1996). This hypothesis argues that both the UG and the entire L1 grammar form the IL initial state. FT/FA suggests that the steady-state IL grammar might converge on the target-language grammar or diverge from that grammar depending on the learner's L1 properties and target-language input.²⁰ This hypothesis will be adopted in the present study (see section 4.2.3, "Research Hypotheses") as it seems to provide the most logical "explanation for the routes which L2 learners take in moving, over time, from no knowledge of the L2 to the eventual mental representations that they construct" (Towell and Hawkins, 1994, p. 132). FT/FA predicts that:

²⁰ Even if there is sufficient input, convergence on the target-language grammar is not guaranteed.

- (1) At the initial state
 - (i) L2 learners are influenced by their L1 in the initial state (see the research relevant to null-subject parameter transfer in section 3.3, below).
 - (ii) L2 learners have access to UG not necessarily via L1 (see research concerning overt pronoun constraint [i.e., Kanno, 1997; Perez-Leroux and Glass, 1999]).
- (2) At the intermediate state(s)
 - (i) The course of acquisition differs from one learner to another; this can be predicted from the assumption that the learners' L1s are a source of their L2 unconscious knowledge (see also the discussion about the variability to be presented below).
 - (ii) The acquisition process may fossilise at this stage, since restructuring does not necessarily take place (see Selinker, 1992).

3. At the final state

Native-like competence at the end state may be attained, though not guaranteed (cf. Birdsong, 1992; Sorace, 1993, 2003; White and Genesee, 1996).

The discussion so far has shown that the parameter approach has opened up various avenues of enquiry regarding the course of second-language acquisition, particularly since the 1980s and '90s. It enabled researchers to examine various issues, including whether adult L2 learners still have access to UG, whether they transfer L1 parametric values to the L2, and whether they can reset L1 parameter values to the values appropriate to L2. Although the debates concerning the broad general questions have by no means been resolved, researchers' interest from the mid-1990s to the present day has shifted towards a closer, more detailed examination of the properties of the IL representation in an effort to understand the nature of the grammars that L2 learners construct. The kinds of questions researchers have been asking are becoming more focused; rather than testing the availability of UG itself in general or the total effect of L1 transfer, L2 researchers test the availability of submodules of UG and the L1 transfer of specific grammatical properties. For example, following one of the developments in the context of the Minimalist Programme (Chomsky, 1995, 1998, 2001), which restricts the differences among languages to functional categories, SLA research in the area of morphosyntax has focused on L2 learners' ability to acquire functional features such as tense, person, number, case, and so on. Chomsky (1989, 1995, 1998, 2000) argues that lexical

categories such as verbs, nouns, and so on are derived from an invariant universal vocabulary and that therefore only functional categories (e.g., person, number, gender) are parameterised; as such, “not all languages make use of all the possible features, and languages differ from one another in exactly how these features are bundled together into individual lexical items and functional categories” (Ionin, 2013, p. 526). (See Hegarty [2005] for more details on how features can be assembled or bundled together into functional categories or lexical items differently from language to language.) Accordingly, acquiring an L1 is determined by selecting the relevant features and assembling them into lexical items and combining them into clauses via a set of universal computational devices such as *merge*, *agree*, *move*, and so on.

The claims accounting for the failure of or persistent problems in L2 acquisition of features fall into two classes, attributing it either to a permanent deficit in the syntactic representation itself (e.g., Hawkins and Chan, 1997; Hawkins, 1998; Hawkins, 2001b; Tsimpli, 2003; Hawkins and Hattori, 2006; Tsimpli and Dimitrakopoulou, 2007) or to mapping problems between syntax and morphology (e.g., Lardiere, 2000, 2007, 2008, 2009; Prévost and White, 2000; Montrul, Foote, and Perpiñán, 2008; Slabakova, 2009; White, 2010). Advocates of the first position believe that acquiring new L2 uninterpretable features that were not selected by the learner’s L1 is impossible once the critical period has passed.²¹ This view is referred to in the literature as the *failure functional features hypothesis* or the *interpretability hypothesis*. The second position about persistent problems in L2 acquisition of features comprises two separate hypotheses: the missing surface inflection hypothesis (Prévost and White, 2000) and the feature reassembly hypothesis (Lardiere, 2008, 2009). Advocates of the missing surface inflection hypothesis argue that “L2 learners have unconscious knowledge of the functional projections and features underlying tense and agreement” (Prévost and White, 2000, p. 113).

²¹ Uninterpretable features are those which play no role in the semantic interpretation of syntactic expressions (i.e., person and number in finite T constituents), whereas interpretable features do play a role in semantic interpretation (i.e., tense and aspect). For a detailed discussion about the distinction between interpretable and uninterpretable features, see Chomsky (1998), Pesetsky and Torrego (2001, 2007), and Radford (2004).

Nevertheless, L2 learners sometimes have difficulty retrieving forms in that the optionality or variability in their use of functional categories or features associated with tense and agreement is attributable to difficulties with the overt realisation of the inflectional morphemes.

The feature reassembly hypothesis presupposes that the main source of difficulty and/or variability in L2A arises when features already existing in the L1 need to be morphologically reassembled into new functional categories and lexical items in the L2 – namely, when the L1 and L2 encode the same feature differently. According to Lardiere (2008, p. 235), this is not just a matter of whether features are still available for selection from a universal inventory and learners select them. Rather, she says, “we need to consider how they are assembled or bundled together into lexical items (or functional categories), and then we must further consider the particular language-specific conditions under which they are phonologically realized” (p. 235).

It can be inferred from the previous discussion that optionality/variability is a central feature of the adult L2 learner’s IL system, even that of a very advanced learner (Sorace, 2003). Thus, one can conclude that the L2 learner’s language system may differ from that of the native speaker in the degree to which target grammatical forms are employed. It is noteworthy at this point that optionality in L2A is not limited to the area of inflectional morphology; its existence in other acquisitional areas including syntax (e.g., word order) has frequently been reported.

The L1 is still seen as one of the main potential contributors to a learner’s IL system. An example to illustrate this characteristic of developing L2 grammars comes from studies conducted to investigate null subject parameter transfer (see section 3.3). Although the role of L1 in null subjects in L2A is still not entirely clear and has been the topic of much debate over the last few decades, there seems to be evidence that L2 learners initially transfer their L1 setting of the parameter to the L2. Moreover, there is much evidence that individual L2 learners (especially at lower levels of proficiency) frequently show variability with respect to their use of pronominal subjects in that a learner sometimes uses the overt pronominal form and sometimes uses the null form of the pronoun. A number of studies conducted to investigate

whether L2 learners transfer their L1 parameters will be critically reviewed in the following section.

3.3 Review of L2 Null Subject Parameter Studies

One of the most intensively studied phenomena within L2 acquisition literature on UG access is the null subject parameter. It has attracted the attention of many SLA researchers over the years (e.g., Al-Kasey and Pérez-Leroux 1998; Clahsen and Hong 1995; Davies 1996; Hilles 1986, 1991; Judy 2011; Judy and Rothman 2010; LaFond 2001; Liceras 1989; Liceras and Díaz 1999; Meisel 1991; Orfitelli and Grüter 2013; Phinney 1987; Rothman 2008; Sauter 2002; Tsimpli and Roussou 1991; Vainikka and Young-Scholten 1994, 1996; White 1985, 1986; Yuan 1997). This is because this grammatical property “motivates a series of theoretical claims about the possible interaction between language typology, UG, and the language acquisition process” (Zyzik 2008, p. 65); namely, it provides a suitable testing ground enabling linguists to examine various issues related to L2A (e.g., UG access, L1 transfer, parameter resetting).

More specifically, investigating the knowledge of the overt/null pronominal subjects and their assumed syntactic and morphosyntactic clustered properties allows insight into learners’ knowledge of finiteness and functional features. This has resulted in a large number of studies investigating the null subject parameter in L2 learners’ ILs. Below, some of these studies are critically reviewed, focusing exclusively on those that concentrate on the L2A of [- null subject] languages and not on the L2A of [+ null subject] languages. This is not only because the present study looks at the L2A of [- null subject] English but also because the majority of the existing L2 studies examining L2A of [+ null subject] languages, such as Spanish, do not address interfaces between discourse and syntax in learners’ IL grammars (for exceptions see Margaza and Bel, 2009; Quesada and Blackwell, 2009; and Rothman and Iverson, 2007, Rothman 2008; Rothman 2009). Without considering the conditions under which subjects must be overt, can be null, or must be null in languages involved in studies, it is impossible to tell whether L2 learners of [+ null subject] languages have reset the L1 parameter value. This is important because

those studies produced conflicting and inconclusive conclusions (i.e., Al-Kasey and Pérez-Leroux 1998; Liceras 1989; Liceras and Díaz 1999; Pérez-Leroux and Glass 1999).

3.3.1 Early Studies

White (1985) was the first study to investigate whether L1 setting of the null subject parameter transfers. She compared the performance of 54 Spanish-speaking native speakers and 19 French-speaking native speakers L2 learners of English by level of proficiency; the learners were classified on the basis of their scores on the Michigan Placement Test from level one to level five. The participants were tested by means of a grammaticality judgment task in which they had to choose one of two responses—*correct* or *incorrect*—and to indicate whether they were *sure* or *unsure* of their judgment and if they chose *incorrect*, they were also asked to try to indicate where a sentence was ungrammatical. The results indicated that the Spanish-speaking learners accepted significantly more ungrammatical English sentences with null subjects than did the French-speaking learners. This contrast was attributed to transfer from L1. Accordingly, White concluded that L2 learners do carry their L1 value of the null subject parameter over into the L2; however, White argued that the gradual improvement in level, as measured by rejecting the ungrammatical sentences with null subjects, indicated that learners eventually can reset the parameter.²²

Similar to her 1985 study, White (1986) compared the performance of two groups of speaking adult L2 learners of English by level of proficiency: one group included 34 adult native speakers of Spanish and Italian [+ null subject], and the

²² Note if parameters can be reset, a sudden jump in the learner's performance would be expected whereby learners, especially those at the advanced level of proficiency, would perform completely native-like. However, it is difficult to tell whether the advanced learners in White's study were really at the end-state of the developmental process; they might have been somewhere at in a lower advanced level of proficiency, yet not at the end-state; that is, they were not yet at the state at which learners are expected to converge on the target-language grammar (see Lardiere, 1998a, 1998b; Long, 2003; and White, 2003 for relevant discussion).

second group involved 37 native speakers of French [- null subject]. The participants were grouped into low-intermediate and high-intermediate levels according to the University of McGill's Placement Test. The participants were tested by means of a grammaticality judgment task in which they had to choose one of three responses—*correct*, *incorrect*, and *not sure*—and were asked to indicate where a sentence was ungrammatical if they chose *incorrect*. White found that the Spanish and Italian participants were much less accurate than the French participants at rejecting the ungrammatical English sentences with null subjects. She concluded that L2 learners transfer their L1 value of the null subject parameter over into the L2, at least initially; however, they can reset the L1 parameter value to the L2 appropriate value with time and increased proficiency.

These studies have often been cited as evidence for null subject transfer and parameter resetting, but they have the following problems. In White's 1985 study:

- i. The French participants' levels of proficiency ranged from one to five (from beginner to advanced). Despite the fact that language proficiency tests do not reflect learners' real language ability, in the sense that there is a considerable amount of uncertainty surrounding the results obtained from any language proficiency measurement (see the discussion in subsection 4.3.5 and in Chapter 5), White grouped and analysed the results of level one with level two, and did the same for level four with five, without giving an acceptable justification. She did so because there were few L1 French participants in levels one and five. Therefore, it could be argued that the French participants, as a subgroup, accepted significantly fewer ungrammatical English sentences with null subjects than did the Spanish-speaking learners as a result of this procedure.
- ii. The test contained few (only four) ungrammatical sentences with referential missing subjects. Yet White marked one sentence as ungrammatical that might be acceptable in colloquial English: John is greedy. Eats like a pig.
- iii. White used a two-point scale for judging the test sentences—correct and incorrect. Many problems and pitfalls are associated with such a scale; see the detailed discussion in subsection 4.3.2.1.2 below.

- iv. The participants were given as much time as they needed to finish the task, which contained only 31 sentences. White mentioned that most of them completed the task in fewer than thirty minutes. However, she admitted that this procedure allowed the participants to change their minds a number of times. This may indicate that the results that emerged do not reflect the learners' IL competences.
- v. Apart from these methodological pitfalls, the data was analysed in a way that makes one remain sceptical about the conclusions drawn from White's results. In order to avoid repetition, this point will be discussed in detail in Chapter 5, where the results from the present study are presented and discussed.

White's (1986) study suffers from nearly the same methodological problems associated with her study conducted in 1985. The test included only five ungrammatical sentences with referential missing subjects; a three-point scale for judging the test sentences—*correct*, *incorrect* and *not sure*—was used, which is problematic; and the results, again, were analysed in a way that makes one sceptical of the conclusions drawn from the study. Moreover, the participants were not well matched. The French participants, as Canadians, probably had been exposed to English. In contrast, the Spanish participants had not lived in any English-speaking communities before they came to Canada.

Let's look at a few more weaknesses of the early studies. To examine whether all L2 learners, regardless of their native language, are able to reset their L1 value of the null subject parameter if transfer exists, Phinney (1987) investigated the use of missing and overt subject pronouns on written compositions of beginning and intermediate adult Spanish learners of English and adult English learners of Spanish as an L2. Phinney argued that both groups of learners were accurate in terms of target-like use of referential subject pronouns. On the one hand, the English-speaking learners of Spanish omitted referential subjects at percentage rates of 83% and 65%, respectively, for the beginning group and the intermediate group. The Spanish-speaking learners of English, on the other hand, omitted the referential subjects at percentage rates of 13% and 6%, respectively, for the beginning group and the intermediate group. Phinney interpreted these results as evidence that the L1 value of the parameter can be reset. However, based on the learners'

performance on non-referential subject pronouns, where the L1 Spanish speakers groups were not as accurate in using the impersonal pronouns in accordance with the target L2 English as the L1 English-speaking groups in accordance with the target L2 Spanish, Phinney argued that the English-speaking learners of Spanish reset the parameter easily compared to the Spanish-speaking learners of English. She related her findings to the theory of markedness, following Hyams (1986) in assuming that the [+ null subject] parameter value is the unmarked setting and [- pro-drop] is the marked setting.²³

However, there are two main methodological problems concerning Phinney's study. First, the two groups wrote their compositions under two different conditions. The composition was part of a regular class exercise for the L1 English speakers and as part of an exam for the Spanish speakers. The exam condition possibly evoked the learners' explicit knowledge about grammar. In other words, they might have focused on the grammatical accuracy as a result of being evaluated. Second, Phinney did not measure the proficiency levels of her participants; she assumed that they were at the higher beginner to lower-intermediate level based on the levels of the language courses they were enrolled in at the two universities. Defining language proficiency is not a simple task (see the discussion in section 4.3.5). Therefore, one must wonder whether Phinney's participants were assigned to their correct levels of proficiency. It seems to me that they were not. Evidence to support this claim comes from Phinney's results. If learners were grouped into their accurate levels of proficiency, why did the beginning group of L1 English speakers perform better with null referential subjects compared to the intermediate group? This may indicate that there is a problem with Phinney's interpretation of her results, i.e., that English L2 learners of Spanish managed to reset the null subject parameter early in IL

²³ This theory has lost its credibility since the introduction of the Minimalist approach to UG by Chomsky in (1993), which reduces the role of UG, compared to how it was viewed in the eighties and early nineties. This approach, unlike Principles and Parameters theory, assumes that the options are not specified by UG; instead they arise when UG does not specify a value. The acquirer's task is to learn, based on the linguistic input he or she receives, which settings are appropriate for the language being acquired for each variant grammatical property. For detailed discussion about the Minimalist approach to language acquisition, see Section 2.3.1.

development. These results are inconsistent with Phinney's findings. The learners' performances suggest that they did not acquire the pragmatic and discourse rules that govern the distributions of null subjects in Spanish; therefore, one can conclude that they did not reset their L1 value of the null subject parameter as Phinney claims.

Two additional studies illustrate early problems in research relevant to the null subject parameter in adult L2A. Firstly, using 13 intermediate- and upper-intermediate-level Greek learners of English, Tsimpli and Roussou (1991) tested whether L2 learners transfer the L1 value of the null subject parameter into L2 and whether parameter resetting occurs. Greek is a [+ null subject] language. The participants were tested by means of a corrective grammatical judgment task and a Greek-English translation task. Tsimpli and Roussou's results with respect to the sentences involving referential null subjects are quite different from the results of White (1985, 1986) discussed above. Tsimpli and Roussou found that both groups of learners (intermediate and upper-intermediate) rejected the sentences with referential null subjects. They also translated the Greek sentences that involved null subjects correctly into English with overt subjects.

These findings suggest that either the property of referential null subjects is not transferred initially or that this property can be reset early in IL development, even before the intermediate level of proficiency. However, based on the learners' poor performance on the other properties assumed to be associated with the null subject parameter (see above), which indicated that the participants had difficulty with sentences involving expletive null subjects and with sentences involving that-trace violations, Tsimpli and Roussou concluded that L1 value of the null subject parameter transfers and cannot be reset to the value appropriate to the L2. They argued against the parameter-resetting approach because they assumed that if parameter resetting takes place in L2A, all the properties associated with that parameter should be acquired simultaneously.

Tsimpli and Roussou's study also presents a number of methodological problems. First, Tsimpli and Roussou did not provide a detailed description of how they measured the proficiency levels of their participants. It seems that they, similar to Phinney, simply depended on grouping their participants according to the information provided by the language courses the participants had completed before

the experiment. This can be inferred from Tsimpli and Roussou's own words, when they say "six of them have already had one year of intensive training in English (intermediate level) and seven subjects have already completed two years of intensive training (post-intermediate level)" (Tsimpli and Roussou 1991, 154). Second, although Tsimpli and Roussou did not describe the scale they used in the grammaticality judgment test to measure the learners' reactions toward the sentences, all the scales used up to this point are problematic in the sense that they fail to provide an accurate picture of IL competence. This is because these scales do not allow sharp lines to be drawn between the learner's certainty, doubt and lack of knowledge reflected in his or her judgements (see the detailed discussion in subsection 4.3.2.1.2). Third, their study involved only 13 participants, who were divided into two groups of proficiency. A larger sample size would have provided a different picture; the larger the sample is, the more accurate the results will be if all the variables affecting the learners' performance are controlled. Fourth, the test contained only five ungrammatical sentences with referential missing subjects, three of which involved copula verbs. Since all the participants had formally learned the target L2, it is likely that they had been taught that *to be* verbs (am, is, are) agree with their subjects in person and number—a factor that might have helped the learners perform well on the test with regard to referential subjects. Furthermore, the participants had to judge and translate short test sentences with referential null subjects restricted to sentence-initial position. These sentences were presented without context. Contextual information is important not only for the interpretation of null subjects but also for the distributions of null subjects among null subject languages.

From this analysis, one can conclude that Greek subjects can be omitted from sentence-initial position only if their references are clear from the context. Hence, one can argue in relation to Tsimpli and Roussou's test sentences that the referents were highly likely brought into focus and, therefore, the learners preferred to use overt referential subject pronouns. This factor might affect the generalisability of Tsimpli and Roussou's results.

Finally, Hilles (1991) analysed the naturalistic acquisition data of two adult L1 Spanish speakers who were L2 learners of English.²⁴ The data came from Cazden, Cancino, Rosansky and Schumann's 1975 study. The purpose of Hilles's study was to investigate the development of overt subjects and agreement inflection on verbs. Hilles's data showed no indication of development with respect to either phenomenon during the course of the nine months of data collection, a finding which also suggests that adult learners cannot reset the L1 parameter value.

However, although conducting a longitudinal study is very useful in reflecting on the nature of IL competence, using data from only two learners to draw conclusions about the process of IL development casts some doubt on the generalizability of extrapolations made from Hilles's findings.

3.3.2 Recent studies

Sauter's (2002) longitudinal study was ambitious in its breadth, with participants from different [+ null subject] languages learning different [- null subject] languages. Nine adult L1 Spanish (Argentinian and Uruguayan) and Italian speaking learners of Swedish, German and English participated in this study. Sauter investigated their natural spoken production data, which comes from the European Science Foundation corpus, in search of evidence for L1 transfer of the null subject parameter value into the ILs and of evidence for parameter resetting. The data cover the early stages of L2A up to about two years of subsequent development. Sauter found that all learners continued to omit referential subjects throughout the data collection. Therefore, she

²⁴ In this study, Hilles also investigated data from two other groups of L1 Spanish speakers L2 learners of English: two young children and two young adolescents. However, because child L2A is different from adult L2A (cf. Haznedar and Gavrusseva 2008; Lakshmanan 1994) and because it is not clear if the CPH applies to children at age 10 (see section 2.2), only the data from the adult L2 learners will be discussed here. It should be mentioned that the results of the two children and one of the young adolescents were exactly the opposite of the results obtained from the adults, presented below.

concluded that the parameter L1 value cannot be reset to the target L2 value. Sauter went so far as to assert the following:

There are no innovations in the relevant IL properties of any individual learner which characterise a new stage in his or her IL development. In fact, the term development may not be appropriate for the IL data that were studied: the absence of new phenomena suggests a lack of development in the properties under investigation [null subject, subject inversion, and that-trace effects]. (142)

However, it seems that there are indeed developmental changes in the proportions of missing subjects; this is inferred not only from the learners' individual data but also from Sauter's own words, for example, when she states that "Nora's [subject omissions] never get higher than 10% after recording viii" (101) and "Tino's subject omissions stabilise around 10% from the fifth recording onwards" (112), although he omitted 100% of referential subjects in the second recording. This progressive improvement may suggest that the learners were on their way to resetting the parameter, as IL restructuring occurs with time and increased proficiency. It can be argued, therefore, that the learners were not at the final stage of L2 development.

Moreover, Sauter stated that "the individual proportions of subject omissions vary strikingly among the learners" (113). This leads to the argument that individual differences existed among the participants at the beginning of the experiment. From the biodata of the individual learners, it can be concluded that the learners were not quite well matched in terms of age of first exposure to the target language (ranging from 20 to 39), frequency of contact with native speakers, proficiency levels at the beginning of the recordings (ranging from no knowledge to elementary), other languages moderately spoken, formal tuition in the target language, and L1 level of education. In addition, although Sauter's study is based on longitudinal data, she has no data of her own; her data comes from the existing published European Science Foundation Corpus, which includes spontaneous data collected from 40 adult immigrant workers. The question is, what are the implemented criteria based on which Sauter chose her nine learners out of the 40? When we examine the data published literature, however, we find that evidence supports both positions, i.e., that the parameter can be reset

and that it cannot be reset. Therefore, Sauter's extreme conclusion that the parameter cannot be reset cannot be taken as certain.

Judy and Rothman (2010) studied two participant groups: 21 native English speakers (the control group) and 18 native Spanish-speaking advanced learners of English. All the participants completed two tasks: a Corrective Grammatical Judgment Task and a Context-Matching Interpretation Task. The first task was designed to investigate whether the L2 learners accepted ungrammatical sentences with null subjects in English. They were asked to judge the grammaticality of the given sentences based on a three-point scale—*grammatical*, *ungrammatical* and *ungrammatical, unsure how to fix it*—and asked to fix any errors if they knew how to do so. Otherwise, they were told to choose the third option—*ungrammatical, unsure how to fix it*. The second task was designed to examine whether there were restrictions on the learners' interpretations of overt embedded subject pronouns in English transferred from their L1. It is possible in English for embedded subject pronouns to have quantified or *wh*-antecedents in the main clauses, whereas in Spanish only null pronouns can do so, and the overt embedded subjects cannot receive bound variable interpretation. This phenomenon is referred to by Montalbetti (1984) as the Overt Pronoun Constraint (OPC). The results of these tasks show that while the advanced L1 Spanish learners of English completely performed in a native-like manner with respect to rejecting the ungrammatical sentences with null referential subjects, a significant difference was found between the English native speaker group and the L2 group in the interpretations of English overt subjects in embedded clauses. The learners' interpretations were constrained by the OPC—a property transferred from their L1 grammar. Therefore, Judy and Rothman concluded that these results, particularly the results of the second task, provided evidence suggesting that the L1 parameter value cannot be reset.

Judy and Rothman's study presents two methodological problems. Firstly, they did not measure the proficiency levels of their participants; Judy and Rothman assumed that the participants were advanced learners based on the fact that 14 of the 18 participants were university students, which indicates, according to them, that they passed the TOEFL test and thus suggests that they are advanced speakers of English (see Judy 2011). Secondly, the rating scale they used in the grammaticality

judgement test to measure the learners' reactions towards the sentences—*grammatical, ungrammatical* and *ungrammatical, unsure how to fix it*—is problematic, because it ignores the possibility that a learner might not have any idea all about the answer.

In the most recent study I will review here, Orfitelli and Grüter (2013) investigated whether adult Spanish-speaking learners of English have an underlying [+ prodrop] in their IL grammar. Seventeen learners took part in the study. On the basis of their scores on the Versant English Test, they were grouped into three proficiency levels: basic user, independent user, and proficient user. All participants completed three experimental tasks: an oral production task, a comprehension task, and a grammaticality judgement task (a replication of the task used in White 1986). The results showed that while the learners accepted sentences with null subjects in the grammaticality judgement task at a mean rate of 41%, they performed in a native-like manner in the production and comprehension tasks; that is, they neither produced sentences with null subjects nor accepted declarative interpretations of null subject sentences—the only interpretation permitted was the imperative one for the sentences with null subjects. However, based on the contrast that emerged when the results of the grammaticality judgement task were compared to the results of the production and comprehension tasks, Orfitelli and Grüter concluded that learners' IL grammars did not license referential null subjects. They argued that if referential null subjects transferred from the L1 to the L2, these learners would be expected to have produced sentences with null subjects and allowed declarative interpretations of the null-subject sentences in English. They claimed that the null subject acceptances observed in the grammaticality judgement task had an extragrammatical cause: general processing limitations.

This study suffers from methodological and conceptual problems. One problem that could invalidate its outcome is that Orfitelli and Grüter were unsuccessful at creating appropriate contexts for subject pronominalization in the production task. The majority of subjects used were full determiner phrases rather than pronominal ones (overt or null). Table 3.1 provides a summary of the different subject types used.

Table 3-1. Subject types used in the production task of Orfitelli and Grüter (2013).

	Full DP	Overt subject	Null subject
L2 learners (n=17)	77.88%	21.50%	0.86%

The overuse of full DP subjects could be attributed to the fact that Orfitelli and Grüter asked their participants to choose names for the two characters in each given picture before they started describing them in only one or two sentences.

Therefore, it can be concluded that the contrast between the results of the grammaticality judgement task and the results of the production task in particular, based on which Orfitelli and Grüter concluded that referential null subjects do not transfer in L2A, is invalid. Hence, one should have reservations about their conclusion.

To conclude this section, based on the literature reviewed, the status of pronominal subjects in adult ILs continues to inspire debate. The results obtained from the different studies seem to contradict each other; that is, there is no agreement among researchers as to whether or not L1 values of the null subject parameter transfer in L2A and whether or not the parameter can be reset to the target language value. Therefore, further research is necessary before we can draw conclusions about the status of null subjects in ILs. The aim of the present study is to build on this research; the following chapter further describes the motivation behind conducting the present study and presents in detail the methodology used to answer the research questions.

Chapter 4. Rationale and Experimental Design

4.1 Introduction

The primary goals of this chapter are to present the research questions and hypotheses to be tested in the present study and to describe in detail the methodology employed to answer them. It is organised as follows: Section 4.2 introduces the research questions and hypotheses, which are formulated based on the gap in our knowledge regarding whether null subject parameters transfer and based on the existed theoretical and SLA research presented in Chapters 2 and 3. This section is divided into three subsections that introduce the motivation behind conducting this experiment, the main research question and sub-questions and the research hypotheses. Section 4.3 includes a detailed discussion of the methodology used in the study. This section is subdivided into seven subsections. The first gives a detailed description of the various groups of participants involved in the experiment. Subsection 4.3.2 discusses in detail the elicitation techniques used in the present study to obtain data from the participants. In particular, this subsection discusses the rationale for choosing a GJ task and a translation task to get an impression of interlanguage at various stages of development; it also provides full details about the design of these data-gathering instruments and some of the criteria implemented in their construction to improve their capacity to provide measurements. Subsection 4.3.3 gives a detailed description of how the data elicitation tasks were refined through two pilot tests; in particular, this subsection will describe these two pilot studies along with the different procedures implemented to increase the tasks' reliability and validity. Subsection 4.3.4 considers the specific procedures involved in gathering the data. Subsection 4.3.5 describes how the participants' different L2 proficiency levels were measured and how the procedure was implemented to increase the reliability of the language proficiency test used for grouping the participants according to accurate measures of their level of proficiency. In subsection 4.3.6, the novel procedure and method applied to score the participants' performance on both elicitation tasks will be described. The last subsection illustrates how the collected data were tabulated and analysed; namely, it

addresses the methodological approach – the analytical procedure – adopted in examining and comparing the gathered data as well as describes the statistical software tool used in analysing the quantitative data and the various statistical techniques utilised to make the necessary reports, comparisons, and contrasts.

Finally, it should be mentioned that large parts of these subsections are devoted to discussing the relevant linguistic and non-linguistic factors that might influence the participants' performance in greater detail and to discussing the various measures and techniques employed to properly control the potential effects of these factors on the reliability of the collected data.

4.2 Research Gap, Questions, and Hypotheses

4.2.1 Research Gap and Motivation for the Study

As we have seen, the various studies reviewed in Chapter 3 (White, 1985, 1986; Phinney, 1987; Tsimpli and Roussou, 1991; Hilles, 1991; Sauter, 2002; Judy and Rothman, 2010; Orfitelli and Grüter, 2013) have produced inconclusive, differing, and even conflicting results about whether null subject parameters transfer and whether those parameters can be reset in L2A. The discussion (that is, the critical review) in the previous chapter made it clear that one of the reasons that these issues have remained inconclusive is the result of serious methodological problems associated with these studies that may have affected the reliability of their data. In light of such methodological flaws, one must remain sceptical that any firm conclusions can be drawn from the results yielded by these studies.

This thesis will be concerned with precisely these matters. This empirical study is designed to address some of the previous studies' shortcomings to answer questions about the issue of null subject parameter transfer in adult L2A in a more accurate way by paying considerable attention to the relevant task-related, subject-related, and procedure- and context-related factors that might influence learners' behaviour. The new measures and techniques employed to enhance the validity of tasks and the reliability of collected data include a newly introduced rating scale and

a novel method for scoring learners' performance, which together promise to solve many of the methodological problems commonly associated with GJ tasks. They also include a new method for placing the participants at their accurate levels of proficiency, among many other practices and constraints applied. Note that this is not to claim that this study is free from weaknesses; the limitations associated with this study are discussed in Chapter 6.

Moreover, unlike all of the previous studies, which involved L2 contexts in which native speakers of consistent- or discourse-null-subject languages were learning non-null subject languages and/or vice versa, the present study for reasons of comparison involves L2 contexts in which native speakers of (a) a partial-null subject language, (b) a consistent-null subject language, and (c) a non-null subject language were learning a non-null subject language – namely, L1 Arabic-, Finnish-, and French-speaking L2 learners of English.²⁵ I am not aware of any study that has been done at this point to compare the acquisition of obligatory subject pronouns in English by L1 speakers of a consistent-null subject language (Arabic) with that of L1 speakers of a partial-null subject language (Finnish). This explains the rationale behind choosing these native languages for the purposes of the study. Furthermore, the syntactic restrictions on the realisation of pronominal subjects in the case of Finnish (see the discussion in Chapter 2) provide an interesting testing environment. Based on a comparison of the Finnish participants' performance in the two contexts – with null subjects either allowed or restricted – differences in their performance in these contexts will indicate whether the tendency to allow null subjects is a non-L1-driven developmental phenomenon or a transfer-related phenomenon. Needless to say, comparing the results of the three L1 groups of learners (Arabic, Finnish, and French) not only will provide us with a more accurate picture of the nature of null subject transfer in adult L2A but also may raise unexpected questions about its nature - null subject transfer - that are worth investigating.

²⁵ Refer to Chapter 2 for a detailed discussion of how these types of languages vary cross-linguistically in the expression of pronominal arguments.

4.2.2 Research Questions and Sub-questions

Based on the research gaps already established and discussed extensively in the literature review (Chapter 3) and briefly summarised in the previous section above, the present study addresses the following research questions:

1. Do referential null subjects transfer in L2A?

To answer this main research question in more detail, the following sub-questions need to be addressed:

- i. Do L2 learners exhibit different developmental paths in terms of null subject parameter resetting? In other words, are there differences in performance between L1 French, L1 Finnish, and L1 Arabic learners of the same proficiency level in L2 English with respect to the syntactic property under investigation?
 - ii. Do L2 learners perform consistently across different task types? In other words, if they accept null subjects, do they produce them, and vice versa?
 - iii. Does L2 learners' performance vary from syntactic structure to syntactic structure? In other words, do different grammatical constructions (e.g., complement clauses vs. adverbial clauses) bring about different performances of accepting null subjects and/or dropping them?
2. If null subject parameter settings transfer, are Arabic and Finnish speakers of English able to reset their L1 value of the null subject parameter to the L2 value?
 3. If null subject parameter settings transfer in L2A, what are the mechanisms that L2 learners use to license and identify the null pronoun?

The results will raise other questions, some of which will be addressed in the course of discussion where possible; others will be left for future research.

4.2.3 Research Hypotheses

Based on the stated theoretical framework (Chapter 2) and in light of the work reviewed above in Chapter 3 about null subject parameter transfer in L2A, the following research hypotheses and predictions are formulated:

Hypothesis H1: According to the FT/FA hypothesis, transfer from L1 will take place based on the syntactic similarities and differences between the L1 and the L2 (positive vs. negative transfer).

Prediction P1: Since FT/FA sees L1 syntactic knowledge as the default source of L2 initial-state syntactic structures and the transitory intermediate states of syntactic knowledge, it can be predicted, owing to the differences in the syntactic realisation of pronominal subjects that exist among the languages under investigation, that:

P1A: The French-speaking learners of English will start with the non-pro-drop value, the Finnish-speaking learners of English will start with partial pro-drop value and the Arabic-speaking learners of English will start with pro-drop value.

P1B: At intermediate stages, the French speakers will retain their non-pro-drop pattern throughout; the Finnish speakers will have problems with 1st/2nd person subjects, but they will not accept a 3rd person null subject (or drop a 3rd person pronoun) unless it is controlled by a higher argument; the Arabic speakers will have the most problems. In other words, the French participants will not accept/drop (null) subjects in English because English adopts the same [-pro-drop] parametric value that French does; whereas the Finnish participants will accept/drop a lower number of (null) subjects in English than their Arabic counterparts as a direct consequence of having an L1 that allows null subjects but under more restricted conditions than in Arabic. As a result of restrictions according to which a 3rd referential subject pronoun cannot be null in Finnish unless it is controlled by a higher argument (unlike 1st and 2nd person referential subject pronouns, which can freely be null), it can be predicted that the Finnish participants would accept/drop more (null) subjects in 1st and 2nd person contexts compared to in 3rd person contexts.

P1C: At the final stage of development both the Finnish and Arabic learners will converge on the English L2 pattern [-pro-drop].

Hypothesis H2: According to the FT/FA hypothesis, the L1 grammar will be reflected in the learner's performance in both tasks, the intuitive and the translation tasks.

Prediction P2: Learners will perform consistently across the two task types. In other words, the degree to which the target grammatical forms are employed will not vary from the grammaticality judgment task to the translation task.

Hypothesis H3: Since all of the L1 syntactic differences are expected to be transferred into the L2 grammar, it will be hypothesised here that L1 transfer does not depend on the structural properties of the target language.

Prediction P3: As it is hypothesised that the different L2 syntactic structures would have no effect on the learners' performance, it can be predicted that all the learners, regardless their L1, would perform consistently across the different syntactic structures involved in the present study (adverbial clauses vs. complement clauses) as far as embedded referential subject pronouns are concerned.

Hypothesis H4: As the entire L1 grammar is assumed to provide the initial state for L2 learners based on the FT/FA hypothesis and since there is link between rich verbal agreement morphology and null subject licensing in some languages including Arabic and Finnish, it can be hypothesised that the agreement morphology of English will be the licenser of null subject in the performances of the Finnish- and Arabic-speaking learners.

Prediction P4: If Agreement (AGR) is the licenser of null subjects, sentences without subject-verb agreement will not be accepted compared with sentences with agreement and null subjects, as this means the Finnish- and Arabic-speaking learners have acquired English agreement/projected English AGR.

4.3 Methodology

4.3.1 Participants

This section describes the groups of participants involved in this study. It considers the following questions: (a) Who were they? (b) What were their different proficiency levels? (c) What controls were incorporated into the selection of participants included in the analyses to make them as homogeneous as possible?

This study involved a large number of participants. In addition to a group of seven native speakers (NS) of English who served as controls, the experimental groups consisted of a total of 487 Arabic, Finnish, and French adult and young adult learners of English as a foreign language (EFL) who were recruited from various universities, colleges, and a high school.²⁶ Table 4.1 offers an overview of the participants at the three different stages of the data-collection processes: the two pilot studies and the main experimental study.

Table 4-1. Participants involved in the study

Stages of test design	English native speakers	Arabic EFL learners	French EFL learners	Finnish EFL learners	Number of participants
Pilot Study 1	0	19	0	0	19
Pilot Study 2	7	14	0	0	21
Main Experiment	0	187	134	126	447
Total Number of Participants					487

However, a large number of participants were excluded from the study final analysis based on the following three implemented criteria:

- i. The first criterion was the learners' performance in the experimental tasks – specifically, the degree to which he or she completed the tasks as instructed.

²⁶ The Arabic participants were undergraduate students at either *Taibah* University or *Umm Al-Qura* University, the French participants were undergraduate students at *Aix-Marseille* University or *Bordeaux* University, and the Finnish participants were students at either *Helsinki* University or a high school.

A detailed description of this excluding criterion is found in section 4.3.6, “Scoring: Procedure and Method”.

- ii. The second criterion was the learners’ achievement on the language proficiency test – in particular, whether he or she had a sufficient level of proficiency in English to perform the experimental test and whether his or her score fell into one of the gaps that were put to draw territories between levels of proficiency in order to address the uncertainty problem surrounding the obtained language proficiency scores. A detailed discussion explaining this excluding criterion is in section 4.3.5, “Measuring the L2 Participants’ Levels of Language Proficiency: Challenge and Solution”.
- iii. The third criterion of exclusion was implemented to minimise the number of subject-related variables among the participants that could have a significant impact on their output; this criterion was established to ensure that they would have a great deal in common, being as homogeneous as possible in terms of age, age of first exposure to English, and other factors.²⁷ The main purpose of this procedure was to control the data so that any differences in learners’ performance could be attributed to the variable of interest – L1 transfer. In what follows, I will explain briefly how these subject-related variables were controlled.

Participant homogeneity was thus achieved, to a great extent, by controlling for the following sources of variables among the participants: age at the time of testing, age of first exposure to English, proficiency level, length of studying English in formal setting (i.e., classes), length of residence in an English-speaking country, and prior linguistic training. As a result, most of the included participants were young adult members of academic communities at the time of testing, but none of them were specialising in linguistics. They all had started learning EFL in a formal setting at

²⁷ The individual differences were controlled for based on an information questionnaire that accompanies the experimental tasks – see subsection 4.3.4.2.

some point between the ages of 7 and 12.²⁸ Furthermore, all three L1 learner groups shared the same level of proficiency in English.²⁹ This means that any participants who had linguistic training,³⁰ had acquired or studied English as a second language (ESL),³¹ had lived in an English-speaking country, or did not have a sufficient level of proficiency in English were excluded.³²

Despite the participants' similarities in terms of these shared features, however, they differed significantly with respect to the following factors:

i. Native language

The experiment participants belonged to three different mother tongue backgrounds: Arabic, Finnish, and French. This is the variable of interest whose effects on the participants' performance the researcher wishes to examine.

ii. Knowledge of other native, second, third, or foreign languages

Though it has been documented in the transfer and multi-competence literature that this factor can influence learners' L2 performance, it was not completely

²⁸ All participants who were exposed to English before the age of seven will be removed from the present study to minimize age effects. A detailed discussion about the effects of age of first exposure was presented section 2.2, above.

²⁹ Note that the individual learners within each learner group differ in proficiency, ranging from lower-intermediate to advanced. This fact will be examined later in this section.

³⁰ Cowart (1997) and Schütze (1996) argue against the use of linguists as informants, as their performance is likely be influenced by their linguistic theoretical knowledge.

³¹ The distinction between ESL and EFL is that the former takes place "with considerable access to speakers of the language being learned, whereas learning in a foreign language environment does not" (Gass and Selinker, 2001, p. 5)

³² Participants who had stayed in an English-speaking country for a short period – two months or less – were not excluded.

controlled for, especially in the case of the Finnish participants.³³ This is because in Finland, in addition to Finnish (the language of the majority population), Swedish is spoken as a national language too. The interaction between the two linguistic communities leads in some cases to bilingualism; as a consequence, some of the Finnish participants had acquired Swedish, in addition to Finnish, natively. Because Swedish – like the investigated language, English – is a non-null subject language, all Finnish participants who had acquired Swedish as an additional native language were excluded from the analysis. Those participants who had learned it after age 7 as an L2 or any other language were included in the analysis. Only a few participants stated that they had not learned any other languages apart from their native Finnish language and the investigated language – English. The situations with the Arabic and the French participants were not the same. Very few French participants said that they were bilingual, though a considerable number had attended German, Italian, Spanish, or Dutch classes. The bilingual participants, like their Finnish counterparts, were excluded from the analysis. None of the Arabic participants mentioned that they had learned any other foreign languages. These differences among the three groups of participants could be attributed to two factors: the linguistic diversity within a particular society and foreign language policy in schools. For detailed discussions of these factors, see, among many others, Ringbom (2002) for Finnish, Payne and Almansour (2014) for Arabic, and Costa and Lambert (2009) for French.

iii. Educational contexts: teaching methodology and pedagogical materials

This is another variable that potentially has a significant impact on learners' performance. This is because their performance depends heavily on the nature of the implicit and explicit input they receive, which in turn depends on the teaching methodology and pedagogical materials they have been exposed to in the EFL context. It seems, in fact, that English teaching in both Finland and

³³ For up-to-date discussion about the effect of multi-competence and transfer at the L3/L4 initial state, refer to Flynn, Foley, and Vinnitskaya (2004) Leung (2006, 2007), and Jaensch (2008).

France, even though most young Finns are much more fluent in English than their French counterparts, is quite good compared to what is offered in Saudi Arabia. While a mixture of the communicative and grammar-translation approaches are adopted to teach English in Finnish and French schools, only the grammar-translation method is used in the Saudi educational context. Moreover, while English educational materials and textbooks are designed to meet learners' needs in France and Finland, they are not in Saudi Arabia. For example, English educational materials in Saudi Arabia have no reference to English-speaking cultures (Gray, 2000); this affected the learners' abilities to use English in authentic situations (Syed, 2003) - a problem which is likely to lead to lack of motivation in the learners (Aleid, 2000).³⁴ For relevant information about the status of English language teaching and learning in these countries, see Ringbom (2002) for the Finnish context; Costa and Lambert (2009) and Perez (2006) for the French context; and Grami (2010), Gray (2000), Whitfield and Pollard (1998), and Rahman and Alhaisoni (2014) for the Saudi context. The following figure summarises the participants' shared and variable characteristics:

³⁴ This situation has started to change; in recent years, more importance has been given to the teaching of English in the Saudi educational context. See Rahman and Alhaisoni (2014).

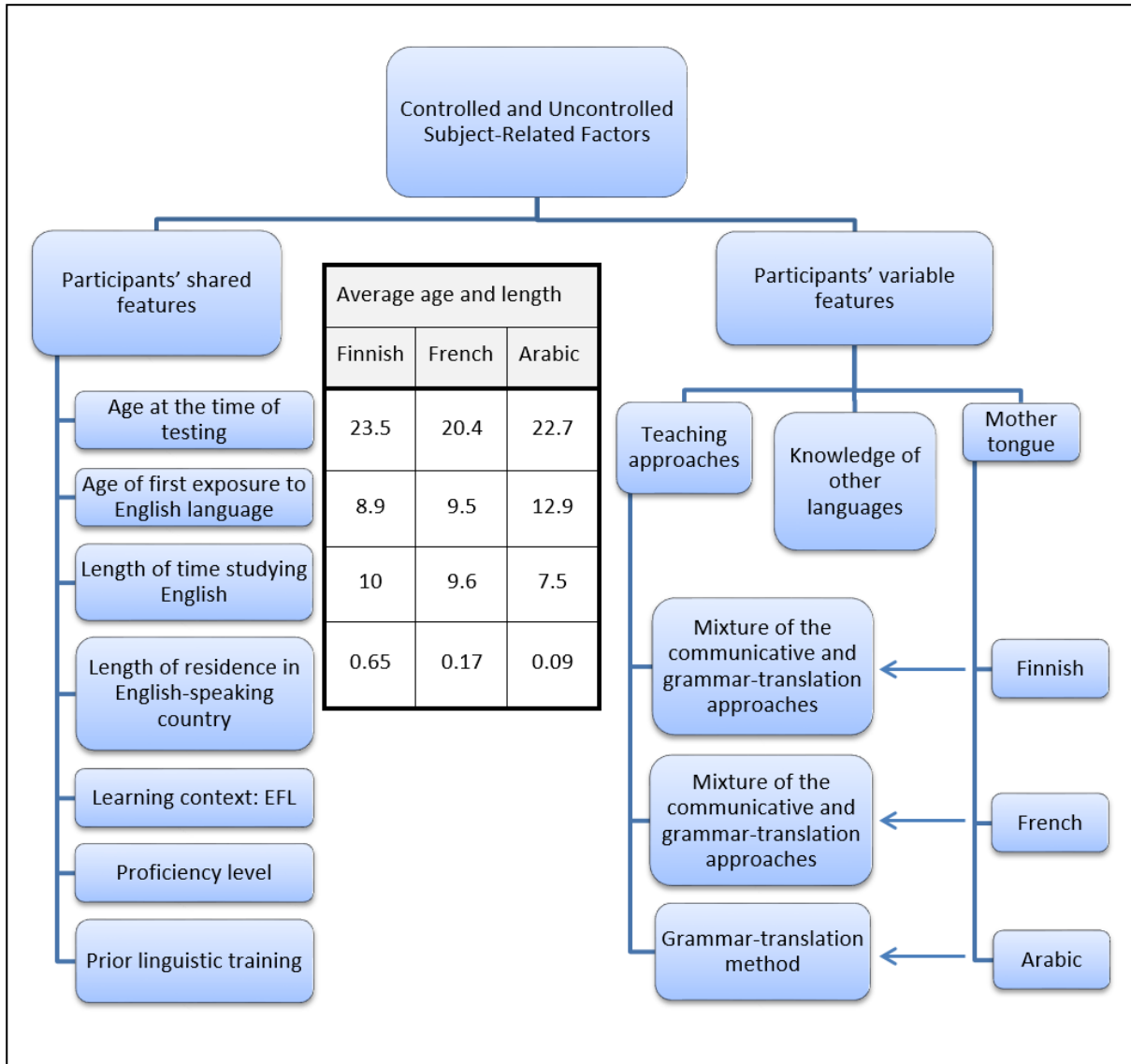


Figure 4-1. The participants' shared and variable characteristics

As a result of these excluding criteria, 263 participants were excluded from the analysis; their percentage was $(263 / 447 * 100 =) 59\%$.³⁵ Hence, the final sample included in the analysis consisted of 184 participants from the three different mother-tongue backgrounds. The participants in each group were further subdivided into three subgroups – lower-intermediate (LI), upper-intermediate (UI), and advanced (ADV) – on the basis of their scores on the proficiency test (see section 4.3.5) to examine how the investigated L2 grammar changed over time at the different

³⁵ 447 was the total number of participants originally in the main experiment; see Table 4.1 above.

developmental stages in relation to the participants' different native languages. Table 4.2 summarises the final total and subtotal numbers of participants in each group and subgroup.

Table 4-2. Participants included in the study

Level of Proficiency	Finnish	French	Arabic	Total Number of Participants
Lower-Intermediate	6	9	16	31
Upper-Intermediate	17	31	14	62
Advanced	53	27	11	91
Subtotal	76	67	41	184

Finally, it should be mentioned that different groups of participants were used for the pilot studies and main study. Only eight of those who participated in the second pilot were included in the main study analysis; see the second pilot (subsection 4.3.3.2). These pilot participants were quite well matched with their counterparts in the main study in terms of their L1 (only Arabic participants), age, age of starting EFL, and so on. These participants will be discussed in section 4.3.3.

4.3.2 Test Instruments: Their Validity, Reliability, and Design

The data examined in this study were obtained from two different elicitation tools: a GJ task and a translation task. The aim of combining two data-gathering procedures was to “maximise the possibility of getting credible findings by cross-validating those findings” (Brown and Rodgers, 2002, p. 243). This mixed methods research is usually referred to as *methodological triangulation* (Dörnyei, 2007).³⁶ Many researchers, including Cohen, Manion, and Morrison (2000, 2007); Weir (2005); Dörnyei (2007); and Clough and Nutbrown (2007), have seen triangulation as an important step to ensure research validity. It offers possible solutions to “reduce the inherent weakness of individual methods by offsetting them by the strength of another, thereby

³⁶ This research approach is also referred to in the literature as multi-methodological research, mixed model studies and mixed methods research; see Creswell, Plano Clark, Gutmann, and Hanson (2003).

maximizing . . . [the] validity of research” (Dörnyei, 2007, p. 43). The following subsections discuss in detail the justifications for using the specific data-collection tools combined in the study. They will also provide full details about the design of these data-gathering instruments and some of the criteria implemented in their construction to improve their validity and reliability.

4.3.2.1 The Grammaticality Judgement Task

4.3.2.1.1 The Validity and Reliability of the GJ Test

As briefly discussed in Chapter 2, from the perspective of Chomsky’s universal grammar, native mastery of a language such as English refers to the learner’s subconscious acquisition of a complex internal system of finite syntactic structures and rules. The system that forms the English grammatical competence of the native speaker, also referred to as the “I-language” (Chomsky, 1986, p. 23), allows the speaker to produce an infinite number of possible utterances and to intuitively reject the impossible ones. Therefore, for researchers to obtain information about the defining characteristics of the syntactic structures that form the grammar of a particular I-language, they must make inferences about what native speakers believe to be grammatical and ungrammatical in the known language.

All native speakers have a built-in ability to automatically and correctly judge which sentences are and are not well formed in their native language. To successfully construct native-like grammar, an L2 learner must acquire similar ability – competence – enabling him or her to determine the well-formedness of sentences in the target language. Therefore, GJ tests have been widely used since the mid-1970s by UG-based L2 researchers such as White (1985, 1986, 1992c), Liceras (1989), Schachter (1989), and Pérez-Leroux and Glass (1997). These and other researchers have used GJ tasks to examine the linguistic competence of L2 learners to find out whether this competence is constrained by the same UG principles and parameters that govern L1 acquisition of grammar and syntax from the initial to the steady-state stages of L2A of grammar.

Recently, however, the use of GJ tests in SLA has been a subject of debate among generative linguists. The serious issue at the centre of this ongoing debate concerns the validity of the research methodology – that is, whether it reflects the L2 learner’s IL. Several researchers have examined the validity of GJ tasks by measuring the extent to which they define the L2 learner’s IL. Some of these researchers, including Bley-Vroman et al. (1988), Han (2000), and Carroll and Meisel (1990), have assumed that GJ tasks directly reflect the learner’s competence. In other words, these linguists suggest that the tests provide a direct window into linguistic competence. Others, such as Gass (1994), White (1989), Cook (1990), and Tremblay (2005), have argued that GJ tasks, like other types of elicitation tasks, cannot directly tap the learner’s IL because competence is an abstraction. This abstract mental knowledge can only be examined indirectly through performance-based evidence, which is often not a perfect reflection of competence.

Such controversy in the literature concerning whether GJ tasks represent the learner’s IL has led to intense debate on the reliability of L2 GJ tests, particularly with regard to whether the “scores on the instrument are free from errors of measurement” (Dörnyei, 2007, p. 110). Various studies, such as the test–retest procedures administered twice to the same participants to address the reliability of GJ tasks (e.g., Ellis, 1990; Gass, 1994; Johnson, Shenkman, Newport, and Medin, 1996), have yielded conflicting results. The results obtained from the correlational analyses of Gass’s judgement data showed that the learners’ judgements over time were consistent, leading the researcher to conclude that GJ tasks provide reliable data for L2 research. Nevertheless, the inconsistent responses to judgements over time have led Ellis and Johnson et al. to argue that GJs do not produce reliable L2 research data. These contradictory findings, along with disagreements about what GJs truly measure, have led another group of linguists, including Birdsong (1989), Lantolf (1990), and Christie and Lantolf (1992), to argue further against the use of GJs in SLA research.

It appears that the controversial status of GJs has emerged as a result of misconceptions about the nature of competence, as described by Chomsky:

In practice, we tend to operate on the assumption, or pretence, that these informant judgements give us “direct evidence” as to the structure of the I-language,

but, of course, this is only a tentative and inexact working hypothesis. . . . In general, informant judgements do not reflect the structure of the language directly. (1986, p. 36)

Given that grammatical competence is abstract and cannot be directly measured, any investigations or observations about the properties of this mental grammar/IL must be inferred through performance data. Such data are necessarily influenced by linguistic and nonlinguistic factors (performance factors) such as the learner's level of proficiency, the amount of time given for the task, the mental state of the individual making the judgement, knowledge about specific language rules (to make use of metalinguistic knowledge rather than unconscious knowledge), and the learner's test-taking strategies, such as guessing (see Schütze, 1996, pp. 98–169).³⁷

Though performance data are not highly accurate in reflecting IL, “some aspects of performance are more revealing than others” (White, 1989, pp. 57–58), depending on what the researcher is investigating. The GJs, for example, despite their drawbacks, can yield performance data that are more representative of a learner's competence than can data gathered by any other tool. This is especially significant in the investigation of a learner's knowledge of certain syntactic structures or rules,

³⁷ Therefore, the research should focus on how to enhance the validity and reliability of GJs for measuring what they have been created to measure. This can be achieved by providing scores that are largely free of errors and that can describe the developing IL and the L2 paths with respect to the native speaker's model. Based on this assertion, some crucial questions must be addressed:

- What are the performance factors that affect L2 grammatical judgement performance? (These answers will help us answer the following question.)
- How do the data converge and diverge across tasks? (Learning this will help us find out the following.)
- On which factor does the learner base his or her judgements? Alternatively, which factor underlines the learner's judgement? (Answers to these questions will guide us to understand the following.)
- Which data reflect the learner's interlanguage the best? (Gaining this knowledge is the ultimate goal of psycholinguistic research.)

such as the null subject parameter and subject–verb agreement, because GJs reveal information about the learner’s syntactic competence. In other words, such data provide information on acceptable and unacceptable sentences, which is one of the defining features of this abstract mental knowledge.

Another obvious advantage that is commonly cited in support of GJ tasks is that they make it possible for experimenters “to investigate aspects of inter-language which may not otherwise be amenable to inspection” (White, 2003, p. 18). GJ tasks accomplish this by placing certain constraints on the learner so that he or she is “forced to make choices within a severely restricted area of his [or her] phonological, lexical or syntactic competence” (Corder, 1973, p. 41, cited in Gass, 1994, p. 306). In this way, the L2 learner cannot avoid considering the specific structure under investigation as he or she may do freely in the unconstrained production data elicitation tasks. This points to another advantage of using GJ tasks in generative linguistics–related research. Because the participant is required only to judge the given sentences and not generate or articulate them, some performance factors such as slips of the tongue, unfinished utterances, speech misinterpretations, and others that may hide some properties of the underlying competence are eliminated.

These advantages explain not only the popularity of GJ tasks among SLA researchers but also the appropriateness of using GJ tasks in this study, despite their drawbacks, to investigate L2 learners’ knowledge of null subject and its related properties (e.g., subject–verb agreement). The next subsection will thoroughly describe the design of the GJ technique used in the present study to obtain data from the L2 learners. It will also discuss the several measures considered to address the common drawbacks of using GJs in SLA briefly mentioned above. Employing these measures, along with the new rating scale, will enhance the validity of GJs and the reliability of the obtained data.

4.3.2.1.2 The Grammaticality Judgement Test Design

The GJ task comprised 70 English sentences. Twenty of these sentences were grammatical and 50 were ungrammatical. Only 38 of the ungrammatical sentences

were used as experimental items. The remaining 12 ungrammatical sentences and the 20 grammatical ones were control and distractor sentences.³⁸

All the test sentences, regardless of their grammaticality or function (experimental, control, or distractor), contained embedded clauses. The embedded clauses were of two types: (a) adverbial of time and (b) finite complement of V. Examples of the test sentences are given below in 1–3:³⁹

1. Experimental sentences:

a. Sentences with embedded null subjects in adverbial clauses:

*He drives whenever goes to work.

b. Sentences with embedded null subjects in complement clauses:

*John told me that found his money.

c. Sentences with missing 3rd-person subject–verb agreement:

*Anne says that John sleep in class.

2. Control sentences

a. Sentences with overt embedded subjects in adverbial clauses:

The children played football until they left.

b. Sentences with overt embedded subjects in complement clauses:

The children think that she went to work.

c. Sentences with appropriate 3rd-person subject-verb agreement:

Sara told me that her aunt often reads the newspaper after dinner.

3. Distractor items:

a. Sentences with null embedded objects:

*The car stopped before John filled with petrol.

b. Sentences with overt embedded objects:

The man says that the glass broke when his son played with it.

³⁸ These 12 sentences will serve as experimental items for another study.

³⁹ See appendix 1 for a complete listing of the test sentences presented in French version of the test.

As can be seen in 1–3, the distractor and control sentences resembled the experimental sentences in syntactic complexity. The grammatical control sentences covered the corresponding acceptable structures of the experimental ones in English. However, the ungrammaticality of all the distractor sentences, unlike the experimental sentences, results in the null realisation of the embedded object pronoun.

The experimental sentences that tested the learning and unlearning of the null subject and mismatched verb agreement were further broken down into different subgroups, each with different tokens. Those meant to test the learning and unlearning of overt subjects in English were divided into two structural types: sentences with verbal-finite-complement or adverbial-of-time embedded clauses. They were then divided into two subgroups, each with different tokens based on the position of the potential referential antecedent of the missing subject pronoun in the lower clause, as follows:

Subgroup 1 comprised 24 sentences with null embedded subject pronouns controlled by higher referential antecedents (pronominal subjects, proper nouns, or noun phrases) within the sentences in the main clause. Of these, 18 were sentences with embedded adverbial clauses while the other six were sentences with embedded complement clauses. This is illustrated by the following examples:

4. a. *The manager can see you when finishes his work.
- b. *The student told me that finished his homework.

It can be seen that from the above examples in (4a, b) that both types of sentential structures share the same verb, *finish*, in the second clauses. This holds true with all other verbs used in the complement clauses in the test; in other words, every verb used in a complement clause is also used in an adverbial clause. The reason behind this is to see whether participants' performance with complement clauses yields results similar to or different from those of the adverbial clauses (see the section 4.3.3, "Piloting and Trying Out the Instrument").

Subgroup 2 comprised four sentences with null embedded subject pronouns in the adverbial clauses controlled by higher antecedents outside the sentences,⁴⁰ such as the following:

5. *I will not leave his office until pays me the money.

The sentences that were intended to test the knowledge of subject-verb agreement were also split into two sets with five tokens each, according to the animacy of the embedded subjects:

Type 1: Missing subject-verb agreement in sentences with animate imbedded subjects:

6. a. *Mary says that John sleep in class.
b. *Bill claimed that the professor always give too many low marks.

Type 2: Missing subject-verb agreement in sentences with inanimate imbedded subjects:

7. a. *Linda explained that the library open late at night.
b. *John though that the cinema often show films in the afternoon.

As can be seen, the embedded subjects and their verbs in some of the sentences in 6 and 7 are separated by an intervening adverb, such as in 6b and 7b but not in 6a and 7a. The purpose of this design was to determine which type of sentences the participants tended to accept or reject more often – that is, to draw a

⁴⁰ Sentences with complement clauses that had sentences outside antecedents were avoided because in most cases these sentences require discourse contexts for the embedded subject to be interrupted to have external antecedents. It was found in the piloting that some participants, when correcting sentences such as that presented in (i), assumed that the missing embedded subject had an external referential outside the whole sentence.

These participants would correct the sentence “(i) *They believed that got better jobs” as “They believed that we got better jobs.”

line between failure, optionality, and native-like attainment in L2A of morpho-syntactic features.⁴¹

So far, it can be observed from the above examples (1–7) that the experimental sentences investigating the acquisition of overt subject pronouns that include null definite 3rd-person embedded subjects are more common than those including null definite 1st- or 2nd-person pronouns in the same syntactic position. There were only six sentences containing 1st- or 2nd-person covert pronouns, while the other 22 sentences included null 3rd-person embedded pronouns. The reason for my interest in investigating the acquisition of 3rd-person subjects is the cross-linguistic variations among the languages under investigation in the obligatory/optional realisation of this particular embedded subject pronoun (see section 2.3).

It is worth mentioning that the number of test sentences, the selected structures with either adverbial or complement clauses, and the number of items/tokens within each type or subtype of experimental sentences were not determined haphazardly. These were composed based on information from previous research and initial empirical data elicited from the results of the pilot studies, as in the examples that follow:

i. The total number of sentences:

To eliminate (or at least reduce) the effect of fatigue on the judgement task, the total number of sentences to be judged was reduced to 70, following Cowan and Hatasa (1994), who argued against using more than 72 sentences.⁴²

⁴¹ Empirical research that has considered the accessibility of the universal inventory of features in adult L2A within the minimalist programme (MP) framework has yielded contradictory results. As a result, a number of hypotheses that account for the failure of or persistent problems in L2A features have emerged, such as, the missing surface inflection hypothesis (Prévost and White, 2000), the representational deficit hypothesis (Hawkins, 2005; Hawkins and Hattori, 2006), and the feature reassembly hypothesis (Lardiere, 2008, 2009).

⁴² It should be mentioned that this is not always the norm in acquisitional studies. Researchers vary in the number of test sentences they give participants. Some studies

ii. The selected sentential structures:

My concentration on the two selected sentential structures (sentences with either complement or adverbial clauses) was grounded in the following facts:

- a. The overt/null realisation of the subject pronouns in embedded clauses varies among the languages under investigation.
- b. The structures of such clauses are neither so simple that a participant is aware of the purpose of the experiment nor so complex that there is a risk that the participant will reject them because of processing difficulties. This claim of relative ease of processing structures of such clauses, although the L2A of complex clauses in English has received little attention,⁴³ can be based on L1 acquisition research showing that such clauses are easily acquired and “need not present serious parsing problems for children” (Bowerman, 1979, p. 294). For example, whereas several L1 acquisitional studies (e.g., Limber, 1973; Bloom, 1991; Diessel and Tomasello, 2001; Diessel, 2004; and Kidd, Lieven, and Tomasello, 2005) have shown that children start to construct complement clauses early, at around age 2, others (Clark, 1970, 1973; Hood and Bloom, 1979; Bloom, Lahey, Hood, Lifter, and Fliess, 1980; Silva, 1991; Diessel, 2004) have noticed that children at age of three are able to make and comprehend sentences with adverbial clauses. These latter studies have also observed that adverbial clauses of time introduced particularly by subordinating conjunctions such as before, after, when, until, and while are among the first to appear in children’s speech,

have been conducted with only 24 sentences (Gass, 1994); others have conducted tests with 282 sentences (Johnson and Newport, 1989). Because of the practicalities of the testing conditions (see section 4.4.4 regarding the test instructions and procedures), I decided not to include more than 70 sentences to avoid fatiguing the participants, even though providing them with longer tests may have increased the reliability of the test results as long as the variables affecting their judgements were controlled.

⁴³ I am not aware of any study that has been done so far to examine the learning of such clauses in particular – unlike, for example, adjectival clauses.

before others kinds of adverbial clauses of condition, contrast, purpose, result, and so on.

iii. The subtotal of items/tokens within each type and subtype of the test sentences:

In the initial pilot study, the participants were far more likely to accept ungrammatical sentences with embedded null subjects in adverbial clauses than in complement clauses; therefore, more sentences with adverbial clauses compared to complement clauses were included (18 vs. 6) (for more details, see section 4.3.3). The remaining 10 experimental sentences were included to test issues related to licensing the null subject in second-language acquisition (SLA). These 10 and the other 12 ungrammatical sentences with null embedded objects served as distractors from the sentences investigating the knowledge of null-subject parameter values as well as from one another. This way, the participants were unable to discern what the study was investigating, increasing the validity of the test. The other 20 grammatical sentences, referred to above as the control items, allowed comparison of the experimental sentences because a rejection of the ungrammatical sentences did not necessarily mean acceptance of the grammatical ones (i.e., successful acquisition).

This task required the participants to judge the test sentences on a 4-point scale: *clearly correct*, *clearly incorrect*, *possibly incorrect*, and *I don't know*. Learners' intuitions in SLA research are reported interchangeably using a variety of terms – *grammatical*, *ungrammatical*, *acceptable*, *unacceptable*, *correct*, *incorrect*, *good*, *bad*, etc. Though there are theoretical distinctions drawn between these terms – see Birdsong (1989) – the rationale for choosing the terms *correct* and *incorrect* to report intuitions in this study was the likelihood that L2 learners, especially those who are still taught via the use of grammar-translation methods, are used to such terms. The participants were instructed to judge whether the given sentence was acceptable by indicating *clearly correct* or *clearly incorrect*, but only if they were confident in their perception of the sentence. If the participant felt there was an error in the sentence but was not certain, he or she was asked to judge the sentence as *possibly incorrect*. If the participant had no idea about the answer, he or she was instructed to choose *don't know*. (See section 4.3.4.) The scale is illustrated below:

1. Bill wondered where was Mary going shopping.

Clearly correct

Clearly incorrect

I don't know

Possibly incorrect

It should be: _____

The words “It should be” followed by a blank space were provided, following Schütze’s recommendation (1996), to prevent participants from making contextless judgements. In a case where the response was clearly or possibly incorrect, the participant was asked to underline the perceived or possible errors in the sentence and provide a correction in the given space.

This procedure enhanced the reliability of the gathered data because it allowed the researcher to consider the correct unexpected responses, such when a respondent rejected a sentence based on a reason that was not actually related to its grammatical incorrectness. In this way, the possibilities of random and careless responses were minimised.⁴⁴ Besides helping to address these cases,⁴⁵ the double-check procedure (making judgements and providing corrections where required as instructed) assured the researcher that the data provided a relatively accurate reflection the learner’s IL if the other performance factors were controlled for.

The 4-point scale in its new format presents several advantages compared to many other rating scales that have been used in linguistic or related research (for a detailed discussion about the different rating scales, see Schütze [1996] and Sorace [1996]). The 4-point scale attends to some of the inherent problems and limitations commonly associated with these kinds of rating scales. Before discussing the advantages, we will consider some of the problems associated with the other rating

⁴⁴ This is because uncooperative participants might not correct all the sentences they mark as ungrammatical.

⁴⁵ See section 4.3.6 for a detailed discussion about the sentences marking criteria.

scales. For the sake of illustration, it is useful to base the critical discussion on a few recent studies.

A binary, 2-point scale (e.g., *grammatical* vs. *ungrammatical*) is usually avoided in SLA research.⁴⁶ This is because L2 learners are usually asked to make judgements during stages of acquisition at which their knowledge about certain elements of the target grammar is incomplete or even totally absent; this status is typically referred to in the literature as “indeterminacy”.

Indeterminacy in a learner’s developing grammar has led some linguists (e.g., Coppetiers, 1987; Bley-Vroman, Felix, and Ioup, 1988) to include a third intermediate option in their scales: usually, the *not sure* selection. The analysis of this third option can be problematic, however. While Coppetiers’s solution is to consider *not sure* responses as meaning “correct”, Bley-Vroman et al. consider these responses as meaning “incorrect”. Both solutions create serious methodological problems that can affect the reliability of the scores. One of these problems concerns equalizing certainty with uncertainty (or doubt) regarding the student’s feelings about the judged sentences. In other words, judging a sentence as *grammatical* or *ungrammatical* reflects a high degree of confidence about it, while judging it with *not sure* reflects much less confidence. Keeping in mind that the goal of psycholinguistic research is to measure IL, this forces us to ask what these *not sure* responses really reveal to us about grammar development. Other groups of researchers who have used 3-point scales have presented another solution to this analysis problem by excluding all of the *not sure* responses from the analysis. This solution, however, even if it may produce more reliable scores than the other two do, still falls short of producing accurate data, raising the same question as with regard to Coppetiers’s and Bley-Vroman’s solutions. What, exactly, do *not sure* responses tell us about grammar development? Furthermore, another serious problem can arise when respondents give too many *not sure* responses.

⁴⁶ Such two-response scales are relatively common in first language syntactic research.

Such problems and pitfalls associated with the 3-point scale have led some linguists (e.g., Schachter and Yip, 1990; Schachter, 1990) to use a 4-point scale ranging from, for example, *clearly correct* to *possibly incorrect*, in agreement with researchers who treat acceptability as a gradient concept (cf. Sorace, 1996; Tremblay, 2005) in the sense that sentence acceptability depends *partly* on the strength of the learner's preference regarding how to say it.⁴⁷ Yet the results analyses gathered using these scales are still problematic. Schachter (1990), for example, counted the *possibly correct* option as if it were the *clearly correct* option and the *possibly incorrect* as if it were the *clearly incorrect* option, allowing for no distinction between the *possibly* and *clearly* categories. Conversely, Schachter and Yip's (1990) results reflect the separateness of these four options. Such unjustified analyses force us to consider the value of including options that cannot tell us anything about the learner's IL since the options are not part of the analysis in any real way, such as in Schachter's study. How do we, then, interpret the two possibly (in)correct options in light of Schachter and Yip's analysis? In addition, what if none of the options applies, such as in the case where the given sentence is beyond the level of the learner, and therefore not part of his or her IL?

The same problems arise with five-point and other multipoint scales used in GJ tasks. In addition, these scales place a greater difficulty on the learner to choose from

⁴⁷ More importantly, acceptability of a sentence depends on other factors, including the sentence's grammaticality, the context in which it is uttered, and whether it is difficult to parse. Thus, it can be said that not every sentence, even if it is well formed, is considered acceptable by all learners (or even by all native speakers of that language), and not every ungrammatical sentence is considered unacceptable by all learners. Note that the notion of acceptability differs from the notion of grammaticality; acceptability refers to the native speaker's or L2 learner's intuitional judgement about a sentence, whereas grammaticality is a theoretical term used by linguists or grammarians to establish whether a sentence conforms to the requirements of the grammar of the given language. Chomsky (1965) also stated that "acceptability is a concept that belongs to the study of performance, whereas grammaticalness belongs to the study of competence. . . . Grammaticalness is only one of many factors that interact to determine acceptability" (p. 11). For more detailed information about the different between these two notions, refer to Haegeman (1994) and Adger (2003).

multiple options in a task that is already highly effortful. Consider the following 7-point scale used in Gass (1994), where she asked her participants first to judge categorically the (un)grammaticality of test sentences and then assess their degree of confidence or doubt regarding each sentence they judged:⁴⁸

-3	-2	-1	0	+1	+2	+3
definitely incorrect			unsure	definitely correct		

It is very possible that some (if not all) L2 learners find it difficult to differentiate between the middle options of Gass's scale – for example, between +2 and +1 or –2 and –1, or even between –1, 0 and +1.

Compared to other scales, however, data obtained from such multipoint scales would produce more complications. Such complications emerge from the same questions raised above with Schachter and Yip's (1990) and Schachter's (1990) studies, but in greater degree: How should a linguist interpret each symbol (number) that indicates more or less acceptability compared to the other ones? More precisely, how is it possible to map territories between all the symbols, especially the ones in the middle? Do these obtained scores reflect the learners' abstract syntactic knowledge? And what can such a score tell us in general about the process of SLA?

Therefore, we can conclude that the results of previous studies discussed here, all designed to reflect L2 grammatical competence, are questionable, affected by a diverse number of factors including the type of measurement scale and response format used. Such a conclusion poses a serious challenge to linguists to find an alternative rating scale that can overcome these methodological problems and pitfalls. The only way to do so is by using a new rating scale that allows sharp lines

⁴⁸ The scope of this section does not allow for further discussion of the problems connected with all of the multipoint scales used in research.

to be drawn between the learner's certainty, doubt, and lack of knowledge reflected in his or her judgements.

The scale used in the present study managed to map the territory between the three possibilities that capture a learner's feelings towards any given sentence. What is unique about the scale compared to others commonly used in grammaticality judgements (nominal, ordinal, interval, and ratio scales) is how the rating scale works and how the data obtained are analysed.⁴⁹ Related to the former point, the rating scale illustrated above is repeated here with an ungrammatical example:⁵⁰

9. John will not marry until finds the right woman.

Clearly correct

Clearly incorrect

I don't know

Possibly incorrect

It should be: _____

As can be seen, unlike in the scales previously used, only one doubtful category (*possibly incorrect*) was used to discriminate between the learner's certainty (both *clearly correct* and *clearly incorrect*) and lack of knowledge (*I don't know*) regarding a sentence. Each of the "possibly incorrect" options has the implied "possibly correct" meaning in addition to its literal meaning. This is because when a learner has doubt that a sentence is incorrect, he or she also has doubt that it is correct, regardless of the degree of doubt (little or great). Though they have the same embedded meanings, the reason behind the preference for including the category *possibly incorrect* in the test, rather than *possibly correct*, is to allow the researcher to ask the learner to correct the possible error he or she believes could render the sentence ungrammatical.

⁴⁹ For a detailed discussion of the variety of different rating scales used in GJ tasks, see Bard Robertson and Sorace (1996) and Schütze (1996, pp. 77–81).

⁵⁰ The latter point is related to how the obtained data will be scored; this will be the focus of subsection 4.3.6, below.

Although I agree to some extent with the researchers who treat grammaticality and acceptability as gradient concepts, I prefer not to assess the L2 learner's degree of confidence or doubt concerning the judged sentences, at least in way used before. This is not only because such procedures increase the burden on the learner. More important, it is because these scales will likely cause the learner to recruit conscious knowledge when reflecting on linguistic rules, especially in contexts where the grammar-translation methods are still taught. I find it difficult to interpret a learner's judgements on sentences from a multipoint confidence scale in a way meaningful enough to increase our knowledge about the process of SLA.⁵¹ In addition, I think it is extremely difficult for an L2 learner to judge sentences on a multipoint scale based on his or her initial feelings, even when asked to do so. I can imagine that the participants in Gass's study discussed above reflected for a moment on which option best described their feelings towards the sentences before they provided their decisions. Therefore, the learners may have consulted their explicit knowledge of prescriptive grammar concerning the syntactic rule under investigation. If this is typical, such confidence scales produce relatively unreliable results and do not reflect the learner's abstract syntactic knowledge. In the end, "we need to be certain that these judgments are predicated on linguistic principles rather than on some other factors" (Goss, Zhang, and Lantolf, 1994, p. 264) so we are able to describe each learner's interlanguage competence.

A very important question arises from the above discussion – that is, whether the scale used in this study can produce reliable and valid data. This question can be answered scientifically only after analysing the data collected via using this scale.

⁵¹ Although Gass (1994) did not provide interpretations, I think the sentences judged as +2 on Gass's rating scale may indicate that learners' grammatical knowledge included information allowing them to accept the structure of the sentences, but they avoided using them that way for different reasons. How do we differentiate these +2 from +1 interpretations, especially if we know that learners are found to "respond with greater certainty and accuracy to nondeviant strings than to deviant strings" (Hedgcock, 1993, p. 3), agreeing with Ellis (1991) and Sutter and Johnson (1990)? Additionally, things become more complicated when we attempt to provide different interpretations for the middle negative symbols -1 and -2.

Nevertheless, the results analyses of the initial and final piloting of the GJ test have provided evidence that judgement data elicited using this 4-point scale are reliable in reflecting the learner's IL. The researcher's ability to exclude some participants from the analysis and distinguish which responses count as appropriate data can support this claim (see the scoring methods section for further details). In addition to these two points, this scale successfully addressed one aspect of the acquiescence threats, which refer to "people who are reluctant to look at the negative side of any issue and are unwilling to provide strong negative responses" (Dörnyei, 2003, p. 13). It was shown that some of the participants whose performance was native-like tended to choose the *possibly incorrect* option and then correct the sentence accurately (see the analysis chapter).

This does not mean that the scale is free from problems, however. Indeed, three limitations have been observed so far:

1. The scale requires proper training before the main test to ensure that the participants have understood the meaning of the category possibly incorrect.
2. It requires the participants to correct the two (clearly and possibly) incorrect options, which increases the burden on the learner.
3. Based on the highly effortful nature of the GJ test and the requirement to correct erroneous responses, the researcher must exclude participants from the analysis to sustain and increase the reliability of the data.

Having described the task, including the test sentences and the type of the rating scale used to report the learners' intuitions, it is important now to describe the measurements that were considered when designing the task – that is, the measurements that control the various factors that might influence grammaticality judgements. These measurements included (a) the participants' different L2 proficiency levels, (b) the time allowed for task completion, and (c) the sequencing of the sentences/items within the test. Each of these is discussed separately below.

i. L2 proficiency levels

Since the same test was administered to each participant regardless of his or her proficiency level, the controls were incorporated into the construction of the test

sentences to ensure that a learner's potential lack of fluency did not hinder his or her ability to complete the task. These controls concerned each sentence's level of complexity, the number of words, and the simplicity of the vocabulary used.

Because "syntactically more complex sentences induce more syntactic processing relative to syntactically simple sentences" (Osterhout, Kim, and Kuperberg, 2012, p. 373),⁵² all the selected sentences were drawn from elementary and intermediate English vocabulary and grammar textbooks.⁵³ The sentences were changed slightly to make them as simple as possible within the context of the selected structures under investigation; for example, the complex verbal structures, such as the subjunctive, perfective tense aspect, were avoided. Only simple tenses were used.

But sentence complexity and length are linked concepts. They both affect grammaticality judgements for reasons related to sentence processing. It has been found that learners will sometimes reject both highly complex sentences and lengthy sentences with simple structures "due to properties of the comprehension process that are independent of grammatical knowledge" (Schütze, 2011, p. 211).⁵⁴ To reduce this possibility and ensure that processing difficulties did not affect the participants' judgements, all the test sentences were controlled not only for the level of complexity but also for the number of words included. Following Dornyei's recommendation (2003) not to exceed 20 words per

⁵² It should be mentioned here that this not always the case at least when it comes to native speakers. Ferreira, Bailey and Ferraro (2002) argue that when comprehending complex sentences, native speakers often use low-level heuristics. This view is known as the 'good-enough' model of sentence processing.

⁵³ Textbooks included *English Vocabulary in Use: Pre-intermediate and Intermediate* (Redman, 2003), *English Grammar in Use* (Murphy, 2010) and *Fundamentals of English Grammar* (Azar and Hagen, 2011), among others.

⁵⁴ See also Schachter and Yip (1990) and Cowan and Hatasa (1994).

sentence (p. 52), the total number of words in each sentence was between 7 and 13.

Another common cause of sentence-processing difficulties among L2 learners is unfamiliarity with some of lexical items used. Murphy (1997) stated that some learners found “a sentence difficult because they could not understand what a specific lexical item meant” (p. 44). Therefore, since it is known that there is a correlation between vocabulary size and reading comprehension (Cameron, 2002; Qian, 2002; Gallego and Llach, 2009), and to avoid lexical items that could be unknown to some of the participants, the vocabulary used in the task was also considered. This is despite the fact that the test sentences were taken from English textbooks designed to suit lower-level learners. Given that vocabulary size varies among learners depending on their proficiency levels (Schmitt, 2000; Schmitt, Jiang, and Grabe, 2011), and driven by the assumption that high-frequency words are acquired earlier than low-frequency words (see Ellis and Beaton, 1993; Hulstijn, 2008), all the lexical items used in the test sentences met the following three criteria:

1. They were among the 2,000 most high-frequency words used by native speakers, according to the National British Corpus.⁵⁵
2. They were also among the 2,000 words most commonly used by L2 English learners, based on the International Corpus of Learner English and the Longman Learners’ Corpus.⁵⁶
3. If verbs, they were high-frequency according to the above criteria. In addition, because the structures used in this study (complement and

⁵⁵ Adolphs and Schmitt (2004) suggest that if a learner masters the 2,000 most frequent words in English, he or she will be able understand around 90–94% of speech and written texts. However, Llach (2011) stresses that the number of words required to understand such a high percentage of the spoken or written context varies depending on the nature of the task to be performed.

⁵⁶ The International Corpus of Learner English is a corpus containing over 3.7 million words of EFL writing by learners from 16 different mother-tongue backgrounds

adverbial clauses) have received no attention in SLA research, all the verbs used, such as want, make, see, look, think, know, hear, watch, say, win, and jump, were among the earliest learned by L1 speakers (Bloom, 1991; Diessel and Tomasello, 2001; Diessel, 2004).

These criteria implemented in the selection of the lexical items used in the judgement sentences reflect an assumption that none of these words are culturally or scientifically related. Another criterion was therefore identified that limits the subject of the sentences: the topics are common in everyday English so that the sentences will be comprehensible to all participants.

Even with these criteria, the possibility remained that some learners might be unable to understand certain lexical items. This problem was eliminated by implementing another measure whereby a list of English words with their meanings in the source languages was prepared following the pilot studies and given to the participants prior to the test (see appendix 2 and the final pilot study, described in subsection 4.3.3.2).

It is important to mention that certain adverbial conjunctions, such as *since* and *after*, were avoided because such conjunctions in some languages (including Finnish) are prepositions that cannot take clausal complements (Holmberg, p.c.). All the adverbial conjunctions of time used in the study (*when*, *while*, *whenever*, *before*, and *until*), in all the languages under investigation, were prepositions that could take clausal complements.

Because of the criteria implemented in the construction of the test sentences and the fact that the minimum level of proficiency required for a learner to perform the task was lower-intermediate on the Oxford Placement Test, the effects of linguistic and nonlinguistic factors that could have decreased the reliability of the

(Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Tswana, and Turkish).

data were minimised.⁵⁷ More important, these measures gave the researcher confidence that when a participant accepted or rejected a sentence, it was based on his or her knowledge regarding the L2 syntax and was reflective of his or her syntactic competence, rather than being based on his or her inability to understand the sentence.

ii. Time allowed for the completion of the task

Hopkins, Stanley, and Hopkins (1990) argued that results vary among learners depending on different factors, including the time given to finish the task. For this reason, Schütze (1996) discussed the importance of determining a limited amount of time during which participants could provide their grammaticality judgements. He mentioned two advantages of this procedure. First, he suggested that including a time restriction for responses could make it difficult for the participant to evoke his or her explicit knowledge about grammar. Second, it could minimise the possibility that the participant might become aware of the researcher's experimental purpose. In addition, Tremblay (2005) mentioned that time restrictions make it impossible for the participant to go back and edit his or her initial response to the sentence.

The participants in this study therefore had a set amount of time in which to provide responses to test sentences. They were required to finish the task in less than 35 minutes. This limited time was restricted based on information obtained from the pilot study. For additional details on how this specific time was determined, please see the final pilot study in section 4.3.3.

iii. Sequencing of the sentences within the test

The sentences in the test were randomised. Moreover, the six ungrammatical sentences with embedded null subjects in complement clauses were placed separately at the end of the test, beginning with sentence 56. The rationale

⁵⁷ The pilot studies showed that learners who scored lower-intermediate on the proficiency test had a sufficient level of proficiency in English to perform this and the translation tasks successfully. For more details, see section 4.3.3.

behind this additional procedure was to prevent the participants from learning the purpose of the study. Following the initial pilot study, it was found that the learners rejected ungrammatical sentences with null subjects in complement clauses far more than they did sentences with null subjects in adverbial clauses. Such high rejection of these types of ungrammatical sentences indicated that some learners might have been able to discern what the study was investigating if the sentences had been placed somewhere at the beginning of the test.

Despite some weaknesses related to the use of grammaticality judgements as an elicitation task (discussed in the previous section), it is apparent that its use is justified. This is particularly because criteria can be imposed on the sentences and procedures. Besides these imposed measures and criteria, this study employed several other measures designed to enhance the result reliability. These additional measures will be discussed in section 4.3.3 (the pilot studies), section 4.3.4 (data-collection procedures), and section 4.3.4 (scoring procedure).

4.3.2.2 The Translation Test

4.3.2.2.1 The Validity and Reliability of the Translation Test

This section discusses the rationale for choosing a translation task to gain insight into the state of learners' interlanguage at various stages of their development. It considers the revival of interest in the use of translation as an elicitation task in current cognitive and applied linguistics research in general and in this study in particular.

Recall what was discussed in relation to L1 influence in Chapter 2. Selinker (1996, p. 103) argues that "translation equivalents play an important role in the formation of Interlanguage competence as they are an important strategy for learners as they look across linguistic systems". This suggests, according to Selinker, that the variation in learners' ability to translate is related to the variation in their L2 interlanguage grammar, which develops with time and increased proficiency. Thus, the tacit assumption is that L2 learners, since their L1 is part of their interlanguage

competence, would refer to their mother tongue when translating into a foreign language and hence are expected to transfer some of their L1 syntactic structures into their L2 production, whether written or spoken.

Similarly, modern cognitive and translational studies provide further reason for using translation in examining interlanguage. Leonardi (2010), for example, argues that when learning an L2 “there is a cognitive function which immediately calls for translation into one’s own native language” (p. 26). Fodor (1983) and Anderson (1992) both claim that the processes of reading, writing, speaking, and listening all depend on an innate mental translation that automatically and unconsciously activates once L2 learners are faced with such tasks. Hence, we arrive at the conclusion that the learner’s native language with its syntax is the cognitive basis for translation into a foreign language, as it is almost unnatural for a learner not to use his or her native language in thinking, writing, or speaking another language.

Moreover, the results of the prior pilot study has revealed that L2 learners – more specifically, L1 Saudi Arabic learners – tend to transfer the targeted structure when translating from their mother tongue to the target language, English, especially in the case of lower-proficiency students. A detailed discussion of the pilot studies will be given in section 4.3.3.

Another reason for using translation tests in interlanguage empirical studies is driven by UG access in L2A research. Baker (1993) claimed that translated texts into English from different languages helps to “isolate patterns which occur across the corpus, irrespective of whether the source texts are French, Hebrew or Chinese” (p 245). In other words, she argues that learners’ translated texts display, in addition to L1-related features, universal features (i.e., “features which typically occur in translated text [such as the tendency towards disambiguation] rather than original utterances and which are not the result of interference from specific linguistic systems” [Baker, 1993, p. 243]). For example, Baker argued that potentially ambiguous pronouns are replaced in translated texts by forms which allow more precise identification. Greenfield and Smith (1976) refers to this as the Principle of Informativeness (see Chapter 5 for more detailed discussion).

Therefore, it seems from the above brief discussion that there are some valid reasons for using a translation task as a technique to elicit data for examining research questions related to L1 transfer in SLA research. Yet there are also arguments against its use as a means to test linguistic competence in the target L2/interlanguage.⁵⁸ Källkvist (1998), for instance, observes that learners' performance on translation tests does not reflect their real language ability. Such an observation, however, even though the study investigates lexical learning through translation and free composition tests, can be explained by Heltai's finding (1992) that translation is not an appropriate test for lower-level learners because such learners are not prepared to perform such task. Newson's argument (1998) that translation is not a good language test because it requires training seems to support Heltai's finding. Another reason for not using it as an elicitation task in interlanguage research is the possibility that the learner might recruit his or her explicit/conscious knowledge, especially in contexts where the grammar-translation methods are still taught (see section 4.3.1).⁵⁹ In this study, such concerns have been taken into consideration in designing this task. The following section will describe the translation task and the measures employed and modifications implemented to enhance its validity to deal with problems discussed above.

4.3.2.2 The Translation Test Design

A written-production translation test was used in conjunction with the GJT to provide another window into learners' IL; the same ultimate goal was being pursued as in the GJT – to examine adults' L2 English knowledge of the negative setting of the null-subject parameter in English. In particular, this test was designed to examine whether adult L2 learners of English at different levels of proficiency, regardless of their L1

⁵⁸ It should be mentioned that despite the extensive existing research on translation in second-language teaching, there is very little empirical research examining the validity and reliability of using translation tests to assess interlanguage.

⁵⁹ Richards, Platt, and Platt (1992) define the grammar-translation method as “a method of foreign language or second language teaching which makes use of translation and grammar study as the main teaching and learning activities” (p. 231).

background, are able to provide the target obligatory pronouns in embedded contexts when translating from their own language into the target language. The task required the learners to translate 15 sentences from their L1 (French, Finnish, or Arabic) into the target L2 (English). The total number of sentences to be translated was limited to this number to conserve energy for the other experimental test that followed. Two L1–L2 translation structural types were selected, following those of the GJT: 10 sentences with adverbial-of-time embedded clauses and five sentences with verbal finite complement clauses. The reason the items with adverbial clauses outnumbered the items with complement clauses two to one was that the participants in the pilot studies were found to have a greater command of the target-language (TL) syntax of complement clauses than they did of adverbial clauses (for more details, see section 4.3.3). Examples are shown here:

Type 1. Complex sentences with adverbial-of-time embedded clauses, as in 1a–c

1. a. J'étais très fatigué(e) quand je suis arrivé(e) à la maison. [French]
I was.1SG very tired(f) when I am arrived(f) at the house
"I felt very tired when I got home."
- b. Tunsin itseni väsyneeksi kun pääsin kotiin. [Finnish]
felt.1SG self.1SG tired when got.1SG home
'I felt very tired when I got home.'
- c. شعرت بتعب شديد عندما رجعت إلى المنزل [Arabic]
jaʕartu bitaʕabi ʕadid ʕindama rajaʕatu lil manzil
felt.1SG tired very when got.1SG to home
'I felt very tired when I got home.'

Type 2. Complex sentences with verbal finite complement clauses, as in 2a–c:

2. a. Elle dit qu' elle veut acheter une nouvelle voiture [French]
She says.3SG that she wants.3SG buy a new car
"She says that she wants to buy a new car."
- b. Hän sanoo, että haluaa ostaa uuden auton. [Finnish]
she says.3SG that wants.3SG buy new car
"She says that she wants to buy a new car."
- c. قالت إنها تريد شراء سيارة جديدة [Arabic]
Qal-t ʔanna-haa turid jiraa sʔayara ʕadida
Said(f) that-3SG-F-ACC wants.3SG buy car new
"She said that she wants to buy a new car."

As can be deduced from the above glosses in 1–2, all three sentences in each type not only have the same semantic interpretation but also share the same syntactic structure: two clauses containing two subjects, one in the matrix clause and the other in the embedded clause. However, all of the Arabic and Finnish sentences of both types contain null embedded referential subjects, unlike the French ones, which must have explicit overt referential embedded subject pronouns; otherwise, they would be considered ungrammatical. All of the other sentences in this task follow the same pattern, with special emphasis on the definite 3rd-person embedded subjects rather than on the 1st- or 2nd-person ones in the same syntactic position – the ratio was two to one. Note that every verb used in a complement clause was also used with an adverbial clause. The reason behind this was to make sure that the kind of verb forms that were given to the test participants did not influence the results/what was being investigated, namely if different syntactic structures in the L2 (e.g., complement clauses vs. adverbial clauses) bring about different performance on the overt and/or null realisation of the embedded subject pronouns in L2A of English (see the section 4.3.3, “Piloting and Trying Out the Instrument”). The complete different versions of the test are given in appendices 1a, b, and c for French, Finnish, and Arabic, respectively.

These sentences were designed by applying criteria that were implemented when designing the final version of GJT, discussed in detail in subsection 4.3.2.1.2. Criteria included (a) the participants’ different L2 proficiency levels and (b) the time allowed for task completion. To ensure that a learner’s potential lack of fluency did not hinder his or her ability to complete the task, the structures of the source language sentences were kept as simple as possible to make sure their L2 English equivalents would be simple biclausal sentences. The length of the sentences was also controlled; the total number of words in each L2 sentence was between 7 and 13 words. The L2 vocabulary required to perform the translation included some of the most common words used in everyday social English, selected by applying the same criteria implemented when selecting the lexis used in the GJT. Since the purpose of this study was not to test the participants’ mastery of vocabulary but to test the acquisition of some properties of L2 syntax in relation to proficiency level, each sentence was followed by some suggestions of words that the participant could use

in translating the sentence. This procedure eliminated the possibility that the learner would be unable to find the right L2 words to match words in the source text.

With regard to the amount of time in which the translation had to be produced in English, the participants had to finish this task within 15 minutes.⁶⁰ This meant that they did not have sufficient time to edit or change any sentences after translating them; thus, their translations can be expected to reflect their unconscious interlanguage knowledge.

The order in which this task appeared during testing was taken into consideration as well. Following the first pilot study, in which the translation task directly followed the GJ task, it was noticed that some learners at the advanced level had figured out that there was a missing embedded subject in a considerable number of sentences in the GJ task, which affected their performance in the translation task. Since this study is investigating learners' unconscious knowledge rather than their conscious knowledge, the translation task was administered before the GJ task in both the second pilot version and the final revised version of the tests to avoid or at least minimise the possibility that some of the participants would discover the purpose of the study before completing the translation task. It could be argued that because learners in translation task were given grammatical sentences to translate into the target language that it was almost impossible for them to discover the purpose of the study. However, when it came to the GJ task, which contained both grammatical and ungrammatical sentences, some learners, especially those at the advanced level, may have discovered the purpose of the study based on the errors that rendered some of the sentences ungrammatical.

It can be claimed that although not all weaknesses related to the use of translation as an elicitation task discussed in the previous subsection have been dealt with, its use is justified, particularly after imposing criteria on the sentences and on the procedures. For more details about how the training requirement problem

⁶⁰ For more information on how the time limit for finishing this test was determined, see 4.3.3.2.

mentioned in the previous subsection was resolved, see section 4.3.4, “Data-Collection Procedures.”

4.3.3 Piloting the Instrument

The quality of the data depends on the ability of an instrument to measure what it has been designed for, providing reliable scores that are free of errors to accurately reflect learners’ ILs in a valid way. This obviously raises an important question: how is it possible to assess the reliability and validity of a particular instrument before using it in the actual experiment? The only possible way to do so is by administering it to a small group of participants similar to the ones who will participate in the study. This small-scale trial is referred to as a *pilot study*. It not only provides a variety of useful information about the instrument’s degree of reliability and validity but also provides information benefiting the test development and administration procedures in a number of ways – checking vocabulary difficulty, evaluating the clarity of the instructions, assessing the time required to complete the task, and so on – in order to “iron out the main problems before the major trials” (Alderson, Clapham, and Wall, 1995, p. 74). This process helps to “avoid the loss of any potentially useful, or even irreplaceable data” (Mackey and Gass, 2005, p. 44).⁶¹ Therefore, the testing instrument for the present study was piloted. In fact, two pilot studies were carried out for all of the tasks in the present study. The following subsections will describe the pilot studies along with the different procedures implemented to increase the tasks’ reliability and validity.

4.3.3.1 Initial Piloting

The initial pilot study was designed to serve as a pre-pilot exercise, mainly in order to determine:

⁶¹ The importance of pilot tests is aptly expressed in the following quotation: “If you do not have the resources to pilot-test your questionnaire, don’t do the study” (Sudman and Bradburn, 1983, p. 283, cited in Dörnyei, 2003, p. 64).

- i. whether the selected data-gathering methods were reliable and valid,
- ii. and whether the target groups of participants were able to perform the tasks as intended by the task developer.

As a pilot test 1, this initial version was not as long as the revised pilot tests 2. It required that the participants judge the grammaticality of 50 English sentences and translate 10 Arabic sentences into English. The structures of the sentences used in this version were similar to some extent to the ones used in the final version of tasks described in section 4.3.2. However, they differed in the types of embedded adverbial and complement clauses used. For the purpose of comparison, this initial version of the tests is provided in appendix 3. This initial pilot study was administered to 19 Arabic EFL adult learners (of whom three were beginner, five were lower-intermediate, five were intermediate, four were upper-intermediate, and two were advanced L2ers of English); as an initial pilot study, it did not use native speaker controls.

Since numerous items were changed, added, and deleted, the statistical description of the results of this initial pilot test will not be reported here, as the scope of the thesis does not allow for detailed statistical description and discussion of the pilot results. However, the findings of this initial pilot study are as follows:

Generally speaking, the methods employed are valid in answering the transfer-related research questions and hypotheses. The following trends were identified:

1. Regardless of their level of proficiency, the participants were likely to accept ungrammatical sentences with null embedded subjects.
2. They were far more likely to accept ungrammatical sentences with null embedded subjects in adverbial clauses than in complement clauses.
3. Their target-like performance with respect to overt subjects was much better in the translation task than in the GJ task.
4. Their overall performance suggested that the acquisition of subject-verb agreement does not cluster with the acquisition of overt subjects.

In this phase, however, it was difficult to know whether the results obtained were valid in reflecting the participants' ILs. The participants' performances could have been influenced by the following uncontrolled factors:

1. Understanding of instructions: The participants asked many questions after the instructions were given to them, which indicated that they did not clearly understand them.
2. Complexity level of test items: Some of the participants asked questions about certain sentences that indicated they had problems with understanding vocabulary items, verbal forms/constructions, and/or sentence length, suggesting that not all participants with lower levels of proficiency were able to perform these tasks. This observation was supported by the fact that the three beginner participants began the test but were unable to finish it.
3. Participant discovery of the focus of the tasks: A few participants mentioned that they had discovered the main focus of the study before they finished the tasks. This was because there were many ungrammatical sentences with null embedded subjects in complement clauses. This type of ungrammatical sentences, as they were rejected by most of the participants, led some of them to notice what the test focused on. This discovery seems to have affected their performance on the translation task, which followed the GJ task.
4. Proficiency-level placement: The placement test failed to provide a clear picture of the participants' current levels of proficiency. For example, it was found that some participants placed in the intermediate category by the placement test ended up performing better on the pro-drop test than those in the advanced category did, which indicated that the proficiency test did not

work very well.⁶² As we shall see, the same problem was also noticed in the second pilot.

4.3.3.2 Final Piloting

Based on the highly informative feedback collected from the initial pilot, the testing instrument was revised. All the pitfalls associated with the initial pilot version were dealt with in this new refined version. Since this revised version was intended to be the near-final version of the main tasks, some of the measures that were implemented to address problems concerning the level of complexity of sentences, the number of words, and the simplicity of the vocabulary used will not be discussed in this section, as they have been described in detail in section 4.3.2 (“Test Instruments: Their Validity, Reliability, and Design”). The solution proposed to deal with the risk of misclassifying the participants in terms of proficiency level will not be discussed; it will be the focus of section 4.3.5. Instead, only factors introduced to solve the problem of the lack of clarity of the instructions will be described here.

It was found in the initial pilot that the reliability and validity of the data were affected by the degree of clarity of the instructions; that is, the quality of the obtained data depended on the participants’ ability to comprehend what the tasks required them to do. Therefore, adequately informative instructions were provided in pilot study 2 to avoid any errors in the learners’ performance that could be attributable to task-instruction-related factors (i.e., errors stemming from ambiguity or

⁶² Of course there must be an explanation for their “odd” behaviour that is well worth exploring. Yet that does not alter the fact that the use of the placement test is made more complicated (or, worse, is invalidated) if some of the lower-level learners, sharing the same L1 as the other participants, turn out to perform better, for example, on the pro-drop test than those at the advanced level. This obviously raises the question: what generalisations can then be made about the different developmental stages based on the results of the pro-drop test in such scenario where some of the advanced-level participants perform worse than those in the lower-levels?

misunderstanding of the task instructions).⁶³ To ensure that the participants fully understood the instructions and would apply them appropriately as instructed, the following fixes were implemented:

- i. The test instructions were translated into participants' different L1s – Arabic, Finnish, and French.⁶⁴ Harmer (1991), Alderson (2000), and Hughes (2003) argue that the advantage of the use of L1s in testing L2 acquisition is that it ensures that the participants, regardless of their proficiency levels, have understood all parts, not one part or some parts, of the task instructions. Similarly, Shohamy (1984) argues that “presenting the questions in L1 may be considered more ethical, since the decision maker obtains information on the test taker’s ability to understand the L2 text” (p. 158).
- ii. Certain highly important parts of the instructions were graphically highlighted – written in bold – for the purpose of attracting the participants’ attention.
- iii. The instructions were further illustrated by examples with model answers for the participants to follow.
- iv. Training and practice examples were included in this version to check whether the instructions were understood and would be applied as instructed. The training sentences were in the same format as the test sentences – complex sentences with either adverbial or complement clauses – but they were varied, investigating the knowledge of different syntactic phenomena. For example, the practice examples for training the participants on how to respond to GJ items required them to apply different types of responses: the first sentence was grammatical; the second sentence was ungrammatical, investigating the

⁶³ This sort of effect is not unfamiliar to SLA researchers (cf. Martinez and Godev, 1994; Cowart, 1997, pp. 55–61). It has been ignored by many psycholinguists, however. Schütze (1996) observes that most of the studies he reviewed did not provide adequate instructions. Furthermore, the majority of the studies I reviewed failed to properly check the participants’ understanding of the given instructions, or at least omitted any mention of that.

⁶⁴ The complete different versions of the test are given in appendices 1a, b, and c for French, Finnish, and Arabic, respectively.

acquisition of the definite article (the); and the third sentence was also ungrammatical according to prescriptive grammar but investigated a persistent long-lasting error in the production of adult learners of English – inversion errors in embedded questions – in order to train the participants how to use the rating category of possibly incorrect.⁶⁵ This sentence with an inversion error is given below:

Bill wonders where is Mary going shopping.

Clearly correct

Clearly incorrect

I don't know

Possibly incorrect

It should be: _____

However, because participants were “likely to be unfamiliar with the linguistic concepts that they are supposed to apply in rating the stimuli” (Keller, 1999, p. 118) and in order to make sure the rating terms used would have the same meanings for different participants (Schütze, 1996), prior to the practice they were given precise instructions regarding the rating formats with which they were required to judge the grammaticality of given sentences. They were told that a sentence should be judged as *clearly correct* if they were sure that the sentence was grammatical (written correctly/has no error) in English; a sentence should be judged as *clearly incorrect* if they were sure that the sentence was ungrammatical (has an error); a sentence should be judged as *possibly incorrect* if they thought there was an error but were not certain about it; and a sentence should be judged as *I don't know* if they had no idea about the answer. For each item ticked as *incorrect* or *possibly incorrect*, the participants were instructed to draw a line under the part of the sentence that they believed was wrong and then to correct the mistake in the space provided underneath the sentence.

⁶⁵ Inversion errors in adult SLA have been reported in the developmental psycholinguistics literature (see, for example, Zobl, 1992; Spada and Lightbown, 1999; McDonald, 2000; Lee, 2008).

As explained, the training sentences focused on syntactic phenomena other than the ones under investigation in the present study to avoid the problem known in psychology as the *Clever Hans effect*, “where the experimenter unwittingly cues subjects into producing the correct answer and thereby overestimates their competence” (Gordon, 1998, p. 224). At the same time, another three important steps were taken to minimise the possibility that some of the participants might discover the purpose of the experiment:

- i. The order in which the experimental tests were presented was changed so that the translation test would appear before the GJ task.
- ii. The test was timed to force the participants to go by their initial reaction to the sentences and not think too much about them; I will briefly return to how the test time was determined in the section below.
- iii. As the total number of the sentences increased to 70 in the GJ task and 15 in the translation task, the subtotal of the distractors and control items increased too – see section 4.3.2.

The structures of the sentences and the individual test items in the second version, which turned out to be the final version of the data elicitation tasks used in the main study, have been discussed in section 4.3.2.

The revised version was piloted for a second time to further ascertain whether the tests were reliable and valid. As native speakers are seen as “the only true and reliable source of language data” (Davies, 2008, p. 431, after Ferguson, 1983, p. vii), the GJ task was first administered to seven British English native speakers.⁶⁶ Not surprisingly, this control group performed as expected. They all rejected all the

⁶⁶ Since the nonnative participants were exposed to varieties of English, it might have been better to pilot the study with native speakers of English from a variety of countries (e.g., the United States, Australia, and the UK). However, since no variations, according to the literature presented in Chapter 2, were found among these varieties of English in relation to the obligatory overt realisation of subject pronouns, collecting data from native speakers of these varieties of English was ignored. In fact, there was no direct access to those informants, and the control participants were all native speakers of British English.

ungrammatical sentences, except in three instances (two participants failed to reject ungrammatical sentences with null subjects, and one participant failed to reject one ungrammatical sentence with missing subject–verb agreement). This acceptance was statistically not significant ($3 / 266 * 100 =$ approximately 1.2%),⁶⁷ and after comparing their inaccurate performance with the other similar but rejected sentences, it was clear that this small percentage of errors resulted from sentence misreading. For reasons of comparison with the results of the learners and also because of space limitations, the statistical analysis of the results of the control group will be presented in Chapter 5. It should be mentioned at this point in the discussion, however, that the control group pilot benefited the study in three ways:

- i. One typographical error was found.
- ii. Two errors were found in one sentence, when sentences were designed to contain only one error – The baby often cries when hears loud noise. This was changed to The baby often cries when hears loud noises.
- iii. It was noticed that the native-speaker participants, when correcting errors in sentences, used the gerund form on rare occasions instead of using definite pronouns with some of the embedded adverbial clauses. For example, sometimes they corrected the sentence ‘*The car stopped before hit the child’ as ‘The car stopped before hitting the child.’ Therefore, if the target groups of participants used either form – the gerund or the pronoun as in ‘the car stopped before it hit the child’– in their correction, this was considered a target-like correction. This is because using the gerund form in this particular position indicates that the learner has realised that English cannot have a null pro subject in this syntactic position. Note gerund phrases in English perform all the functions that nouns do (for more information, refer to English grammar textbooks, e.g., Frank 1972).

After dealing with these two minor errors, this updated version was again piloted with a group of 14 Arabic EFL adult learners (of whom two were beginners, five lower-intermediates, three intermediates, three upper intermediates, and one advanced).

⁶⁷ The total number of tokens was 190, which is the number of experimental items multiplied by the number of participants = $38 * 7 = 266$.

Again, unfortunately this pilot did not use other target EFL learners groups – Finnish and French – because there was no direct access to them for reasons related to budget and schedules.⁶⁸ Because the initial pilot pitfalls had been dealt with, the tasks ran smoothly; the participants understood the instructions thoroughly and unambiguously with ease; they asked very few questions after the instructions were given; they successfully applied the instructions as instructed. However, two beginners began the test but were unable to finish it. It seemed that the test was not suitable for the beginner EFL participants; the target groups, therefore, were lower-intermediate, upper-intermediate, and advanced participants.

This second pilot also yielded results that showed the tests were valid and reliable. The results of this pilot test will not be statistically described in this section; as this piloting phase did not result in major revisions, it was the final version of the tests and some of the data obtained were used in the main study.⁶⁹

It should be mentioned that since one of the purposes of the pilot was to determine the time required to finish the test, the participants were told when given the instructions that the test was timed. They were asked to complete the test as instructed as fast as they could without going back to edit any answers they gave, but they were given as much time as they required to complete the tasks. They finished the translation task in 10–19 minutes, the GJ task in 30–38 minutes, and the proficiency test in 14–20 minutes, depending on their level of proficiency. No editing was noticed. The average time was approximately 15 minutes for the translation task, 35 minutes for the GJ task, and 17 minutes for the proficiency test. These averages

⁶⁸ In fact, the initial pilot was put online and sent to some possible participants in France and Finland, but no one completed both tests.

⁶⁹ Based on the participants' comments after they finished the test, only one word that has a culturally related content, "girlfriend," was replaced in this version with a common item, "wife," in the final version of the test.

were set as the time limits by which the participants were required to finish the experimental tasks in the main study.⁷⁰

Finally it should be mentioned that the participants during the second pilot study were asked to circle all the vocabulary items that they did not understand. Following the pilot, an English vocabulary list with the words' meanings in Arabic, French, and Finnish was prepared to be given to the participants in the main study.

4.3.4 Data-Collection Procedures

This section considers the procedures involved in gathering the data. It also discusses the measures that were employed to properly control the potential effects of the relevant procedural and context-related factors – the practical aspect of administering the data collection tools – on the reliability and validity of the gathered data. These factors include controlling the testing environment, minimising the participants' anxiety level, and the amount of time during which the test was taken.⁷¹

4.3.4.1 The Consent Form

Given that the act of collecting data is often thought as an intrusion into participants' private lives (Cohen et al., 2000, 2007; Denscombe, 2007), the participants were initially given a consent form to read. This form provided the participants with general information about the purpose of the present study: to investigate the intuitive knowledge of English speakers, both native and non-native speakers, and to discover what sort of language is natural for most people. It provided them with concise information about the tasks they would be asked to complete – a translation task, a GJ task, and a language proficiency test – and the approximate amount of time

⁷⁰ As the level of interlanguage competence can vary among participants, so should the average time determined for each level of proficiency to complete the task. This was not possible in the present study, however, because the translation test was administered directly following the proficiency test in the same session, which prevented grouping the participants according to their levels of proficiency.

⁷¹ Such procedural factors are usually discussed in the literature under task-related factors.

required to finish them. The participants were informed that there were no foreseeable risks associated with this study. However, they were instructed that they would be free to discontinue the study at any time if they felt discomfort for any reason, and they could request their answers be destroyed.⁷² They also were told about the direct benefits that they would receive: £8 or the equivalent amount in euros as payment for their participation.⁷³ They were assured that all their information and answers would be treated with complete anonymity and confidentiality. Having agreed to take part in this experiment, the participant was asked to sign and date the consent form. If he or she was a young adult and was considered a minor by law, his or her parent, guardian, or legal representative had to give permission for the minor to be included in the research study.⁷⁴ Therefore, in addition to the consent form to be completed by the adult participants, another form was designed to be completed by the minor himself or herself and by his or her parent, guardian, or representative. Both versions of the consent forms are found in appendix 4. However, on every occasion when the minor informed consent was required, it was signed by the minor's teacher or lecturer.⁷⁵ This was because testing on all occasions was carried out in controlled settings – in universities, colleges, or high schools – to minimise the possibility that participants would become distracted when tested and to reduce the variations between the participants by having all of them perform the tasks in similar testing conditions, following Schütze's advice (1996). Nevertheless, the time during

⁷² Only a few of the test participants discontinued the study and requested their answers be destroyed. Cowart (1997) points out that "the informant's state of mind may well change . . . as she [or he] proceeds through the questionnaire. Fatigue, boredom, and response strategies the informant may develop over the course of the experiment can have differing effects on sentences judged at various points in the entire procedure" (p. 94).

⁷³ For cultural reasons, the Saudi participants were not paid for their participation in this study.

⁷⁴ This is because the minimum legal age at which a person can give his or her consent to participate in a study varies across relevant countries. The minimum age in Finland is 18, in France it is 16, and in Saudi Arabia it is not defined. For more information about the participants' ages and the age variable, see section 4.3.1.

⁷⁵ Thanks to Dr Joachim Zemmour and Ms Inna Smirnova.

which the test was performed was not controlled among the participants because the data were collected from many participants on various occasions in different countries.⁷⁶

4.3.4.2 The Personal Information Sheet

After signing the consent form, the participants were asked to answer a brief questionnaire designed to gather personal information such as a participant's native-language(s), age, age of first exposure to English, length of time studying English in formal setting (i.e., classes), length of residence in an English-speaking country, knowledge of other foreign/second language(s), and prior linguistic training (whether he or she had studied linguistics as an academic subject). The complete information sheet is found in appendix 5. This information was necessary to control and statistically examine the significant effects of relevant participant variables on the outcome of the experiment. For more information about this interrelationship, see section 4.3.1 and Chapter 5.

4.3.4.3 The Administration of the Data Elicitation Tools

The participants then were given the testing instrument's experimental tasks (the translation and GJ tasks) along with the proficiency test and the vocabulary list.⁷⁷ They were instructed on how to carry out the different experimental tasks before beginning.

As the experimental tasks administered differ in their nature (intuition vs. production), each task type requires different instructions. Because the translation task was to be completed before the GJ task, that task was explained to the participants first. They were instructed that they would be required to translate 15

⁷⁶ For more details about the possible effect of testing time on L2 learners' performance, the interested reader is referred to Hopkins, Stanley, and Hopkins (1990).

⁷⁷ This list contained words with translations into the participants' own languages, prepared based on the second pilot study. For more information, refer back to subsection 4.3.3.2.

sentences from their native language into English, using in their translation, if they wished, the suggested words that were located under each sentence. If they did not understand the meaning of an English word, they were told to refer to the vocabulary list given. They were provided with two models to follow. They were asked not to go back to change any sentence after they had translated it. They were told that the task was timed and they would have only 15 minutes on this exercise to translate all of the sentences provided.

With regard to the GJ task, the participants were requested to judge the grammaticality of the sentences based on the given rating categories. They were instructed with examples as to what each option meant, as explained in the previous section describing the second pilot study. They were asked to correct the observed or assumed errors in the sentences. To ensure that the participants applied the instructions and rating scale as instructed, they were given training and practice on three examples. They were informed that this test was timed too, and that they were required to finish judging all the items within 35 minutes. It was made clear to them that they would not be allowed to exceed the time limit (i.e., they were urged not take too much time to provide a response to sentences, but to go by their first impression of the sentences, which meant not going back to change or edit earlier responses).

To reduce the participants' anxiety, they were told that the tests were for research purposes to reassure them not to worry about their scores. Translating the test instructions, as recommended by Shohamy (1984) and Stibbard (1998), was also intended to lower the anxiety level of the test takers. Two relevant factors that played important roles in determining the participants' feelings towards the tasks were the vocabulary list that provided the meanings of English words in the participants' L1s and the suggested words included under each sentence to be translated. The participants were instructed to check them (the vocabulary list and the suggested words) just prior to starting the actual experimental activities, which helped them answer the test items without being afraid of not recognising the meanings of words, of not being able to find the right words to use, or of making spelling mistakes.

After the instructions had been read and the participants had been trained on a number of examples, they were allowed to ask questions; however, very few

questions were asked. The participants appeared to understand what was required and therefore felt ready to progress on to the main experimental tasks.⁷⁸ Before they were allowed to do so, however, they were very briefly instructed on how to perform the Oxford Placement Test. It is a multiple-choice test, and all the participants were familiar with this type of test where they just needed to select an answer from a list. They were informed that they had to finish this task within 17 minutes.

They were reminded that they would be interrupted between one task and another to be instructed to move to the next task once the time limit for the previous task had been reached.

4.3.4.4 The Debriefing Form

Once each participant completed all the tasks, he or she was debriefed about the nature and the exact purpose of the present study. The participant was given the debriefing form along with the cash reward for his or her participation in an envelope. The debriefing form is found in appendix 6.

4.3.5 Measuring the L2 Participants' Levels of Language Proficiency: Challenge and Solution

Defining language proficiency (LP) is not a simple task. Several definitions of this construct have been proposed in the literature (see Lantolf, 1988; Hulstijn, 2010, pp. 185–187). All of them reflect the complexity of the various types of knowledge and skills that LP involves and how complex assessing such knowledge and skills is. Thomas (1994), for example, defines language proficiency as “a person’s overall competence and ability to perform in L2” (p. 330, footnote 1). Such a definition, though straightforward, reveals the complexity of assessing LP. To understand the source of this complexity, two fundamental questions can be raised based on Thomas’s definition of LP: (1) What is competence? (2) What is ability?

⁷⁸ The pilot trials – section 4.3.3 – ensured that the participants would completely master the instructions and the tasks would run smoothly.

1. Competence is the speaker's mental knowledge of the language (see Chomsky, 1965). Such knowledge is an abstraction; therefore, it can only be assessed indirectly through performance.
2. Language ability is "the ability to perform language tasks in real life and real time; that is, the ability to convey or understand a content message through the medium of spoken or written language" (Schoonen, 2011, pp. 701–702). Language ability is usually referred to within the UG framework as language performance, which is defined as "the actual use of language in concrete situations" (Chomsky, 1965, p. 4). Looking at L2 learners' data makes it clear that performance is often not a perfect reflection of competence.

These definitions of LP and its closely related notions of competence, performance, and language ability represent the major hurdles to developing a valid test that can yield reliable results for effectively measuring LP. There is a considerable amount of uncertainty surrounding the results obtained from any LP measurement. This uncertainty can be traced back to a number of external and internal variables that influence the assessment of LP in the L2 learner (see Cook, 1996; Skehan, 1989).⁷⁹ In the face of this ambiguity, Messick (1989) claimed that there was no ideal test that could efficiently measure the L2 learner's exact level of proficiency. Though this ineluctable problem threatens a test's validity and reliability, it does not imply that LP tests should be abandoned, as the assessment of LP is profoundly important for many educational and research purposes. In the context of SLA research, for example, various studies have shown that a learner's performance is mediated by his or her level of proficiency (see Chapter 3 and Chapter 5, where the findings of syntactic acquisition studies show a significant impact of LP levels on the process of L2A). To increase LP test validity and reliability, some researchers (e.g., Lado, 1961; Klein-Braley and Smith, 1985) have argued, based on the assumption that LP involves abstract knowledge and many skill components, that the ideal way to assess LP is through multiple tests, each focusing on a single type of L2

⁷⁹ Some of the variables affecting L2 learners' performance, such as native language, current age, age at which L2 education began, multi-competence, kinds of exposure, and others, have been discussed in the previous sections.

knowledge or skill. After combining the tests' subscores, a more complete picture of the learner's level of proficiency can be deduced. This model of testing is known as the "discrete point" approach. Nevertheless, it presents several limitations. These include the inability to test separate types of knowledge (e.g., grammar) or skills (e.g., listening) without involving others (e.g., vocabulary, speaking) and the potential loss of test efficiency as a result of high testing costs and time-consuming administration, performing, scoring, and analysis of results. These weaknesses of the discrete point testing approach gave rise to the integrative test approach. This testing model assumes that a single test can measure a combination of mixed knowledge and skills (both linguistic and nonlinguistic) if constructed to do so. This is because "language processing or use entails the simultaneous engagement of more than one language component (e.g., vocabulary, grammar, gesture) and skill (e.g., listening, speaking)" (Vecchio and Guerrero, 1995, p. 6).⁸⁰

This brief information about the complexity of measuring LP and the difficulties of designing an efficient LP test suggests that choosing an LP test is not a simple task for the SLA researcher. Hulstijn (2010) argues that a researcher must first consider the "study's goal, research questions and theoretical embedding . . . [before he or she] . . . [can] . . . decide which construct of LP, or which LP component(s) should [be featured] as a variable [variables] and how it [they] should be measured . . . [taking into account] the LP test's [or tests'] proved or expected validity" (p. 185–196). This claim is in agreement with Wistner, Sakai, and Abe (2009), who proposed that "researchers need to choose a testing instrument that measures the aspect of proficiency that is related to a particular study" (p. 33). By doing so, the chosen proficiency test can supply more meaningful information about the link between the participants' levels of proficiency and the purpose of the research study; for example, a test meant to measure syntactic competence will be useful for a study on syntax.

⁸⁰ For thorough reviews of both types of language proficiency tests, refer to Thomas (1994, p. 326), Alderson (2000, pp. 206–207), and Hulstijn (2010, p. 188).

The goal, research questions, and theoretical context of the present study revolve around the acquisition of L2 syntax, particularly L2A of English subject pronouns (see sections 4.2 and 4.3). As such, the best test option appears to be a major existing English proficiency test designed to assess L2 learners' grammatical, vocabulary, and semantic knowledge – the Oxford Online Placement Test (OOPT) (free version). This test was chosen for its focus on testing different aspects of grammar⁸¹ and because it is reliable and efficient. (The test's reliability is briefly discussed below along with the scoring procedure.) The test was free of charge and easy to administer after conversion from an online to a pen-and-paper format.⁸² It requires approximately 15–22 minutes to be completed by the participants,⁸³ depending on the test taker's proficiency level,⁸⁴ and consists of a 50-item multiple-choice grammar test with a total score of 50 points (see appendix 7). The participant's level of proficiency is initially identified by his or her total number of correct responses, according to the following user's guide to classifications of proficiency levels:⁸⁵

⁸¹ The OOPT was originally designed to test learners' grammatical knowledge of English to see whether they could benefit from the English-for-academic-purposes courses at the Oxford University Language Centre (see the test website, http://www.lang.ox.ac.uk/courses/tst_placement_english.html).

⁸² This process did not affect the test's validity, as it did not entail changing the style or the nature of the test. It is a multiple-choice test, in which test takers just need to select an answer from a list; the popularity of this test type means that all second-language learners are familiar with it.

⁸³ The required completion time for the participants in the present study was based on information obtained from the pilot study (see section 4.3.3).

⁸⁴ Shortness was an important test feature after weighing its validity, reliability, and efficiency, as the participants needed to conserve energy for the two long experimental tests that followed.

⁸⁵ I adhere to the descriptors in the Oxford Placement Test because such terms (beginner, lower-intermediate, upper-intermediate, advanced) are commonly used in the recent SLA research (refer to, among many others, Carrol and Conklin, 2016; Solon, 2016; Barrios, Jiang and Idsardi, 2016).

Table 4-3. Oxford Online Placement Test: proficiency-levels classifications scale

Level	Proficiency	Score range
Level 1	Complete Beginner	1–3
Level 2	False Beginner	4–10
Level 3	Lower Intermediate	11–20
Level 4	Intermediate	21–30
Level 5	Upper Intermediate	31–40
Level 6	Advanced	41–50

For the present study, however, this six-level LP classification was reclassified to four levels as follows: Level 1, Complete Beginner, was merged into Level 2, False Beginner, to form the Beginner Level, and Level 4, Intermediate, was deleted.⁸⁶ This process necessitated a remapping of the test scores. This new classification resulted in the four proficiency levels presented in Table 4.4.

Table 4-4. Reclassification of Oxford Online Placement Test proficiency-levels scale

Level	Proficiency	Score range
Level 1	Beginner	1–12
Level 2	Lower Intermediate	13–25
Level 3	Upper Intermediate	26–38
Level 4	Advanced	39–50

The reclassified scale provides a clearer picture of the participants' current levels of proficiency. It aided in the assignment of correct levels of proficiency by reducing the problem of uncertainty associated with any language proficiency scores mentioned at the beginning of this section. This claim is illustrated by the notion of *residual uncertainty*. In a language-testing context, this term refers to the reported amount of uncertainty associated with a language proficiency test score (cf. Pollitt, n.d.) – that is, how accurately it reflects a participant's real language ability. For example, in the commercial version of the OOPT, the residual uncertainty is ± 5 points.⁸⁷ This means that, if a participant scores 18 with an uncertainty of 5 units, his or her real score range is $(18 \pm 5 =)$ 13 to 23, leaving the researcher unable to

⁸⁶ I will return to explain why I did so below, after discussing the notion of residual uncertainty.

⁸⁷ For more details, see Pollitt (n.d.).

determine whether the learner’s level is A1 or A2 (see Pollitt, n.d., pp. 9–10). If we assume that the free version of the test, which was used in this study, presents the same amount of residual uncertainty (± 5 scores),⁸⁸ the same problem would arise, especially if the score obtained was very close to the boundary of any proficiency level.⁸⁹ However, the probability that the scores reflect each participant’s real level of proficiency depends on the levelling of the test used to measure it. That is, the real proficiency is more likely to be reflected in a four-level than in a six-level proficiency scale because of the link between the specified residual uncertainty and the difference in score range between levels of proficiency in each classification scale. Table 4.5 illustrates this statistical fact.

Table 4-5. Assessing accuracy: percentage of uncertainty associated with levels of proficiency: a comparison between the original and the reclassified scales

No. of levels in proficiency Scale	Total test mark	Amount of residual uncertainty	No. of boundaries between levels	Total No. of scores leading to uncertainty about level of proficiency	Percentage of uncertainty about the level of proficiency
Six-Level Scale	50	± 5	5	$5 * 5 = 25$	50%
Four-Level Scale	50	± 5	3	$5 * 3 = 15$	30%

This table shows that, with an uncertainty of ± 5 scores, the probability that the participant’s real level of proficiency could be wrongly classified is quite large with the original OOPT levels of proficiency (about 50%). This misclassification probability diminished to 30% with the reclassified OOPT levels of proficiency. Even if the

⁸⁸ I am not aware of any study that has been conducted to determine the residual uncertainty of scores obtained from this test’s free version.

⁸⁹ Although reassessing participants whose scores were close to a boundary appeared to be a solution to avoid the risk of misclassification, it was not possible due to practicalities of the testing conditions. Nevertheless, it has been found that, as a result of the elusive nature of linguistic competence, not only do different language proficiency tests lead to different language classifications, but a single language learner might get different test scores on different occasions (see Ulibarri, Spencer, and Rivas, 1981; Pollitt, n.d.).

misclassification probability was reduced by 20% compared to the original proficiency scale, however, 30% is still relatively large. To deal with this methodological problem (i.e., to avoid the risk of misclassifying the participants into levels of proficiency determined by the residual uncertainty of ± 5), a five-score gap was left between each two levels of proficiency in the classification scale. This is illustrated in Table 4.6:

Table 4-6. Oxford Online Placement Test reclassified levels of proficiency: score range redistributing/remapping

Level of	Score Range	Number of Scores
Level 1: Beginner	1–8	8
Gap 1	9–13	5
Level 2: Lower Intermediate	14–22	9
Gap 2	23–27	5
Level 3: Upper Intermediate	28–36	9
Gap 3	37–41	5
Level 4: Advanced	42–50	9
Total no. of scores:		50 marks

As can be detected from comparing Table 4.6 to Table 4.5, the five scores that drew territorial gaps between levels of proficiency were deducted from the numbers of the score ranges in the reclassified levels of LP according to a particular calculation to ensure that almost all levels of proficiency shared the same number of score ranges by nine marks.⁹⁰ Accordingly, any participant whose score fell into one of these gaps (9–13, 23–27, or 36–40) was excluded from the analysis. By following this criterion, a decision could be made with almost no uncertainty. Hence, the probability of misclassification dropped to zero.⁹¹ This procedure will increase this test's reliability for grouping the participants into their accurate levels of proficiency.

⁹⁰ Except level 1, "Beginner", with a score range of eight marks. Students whose scores fell in this level were excluded from the analysis, as they did not have sufficient levels of proficiency in English to perform the experimental test. "Lower-intermediate" was the threshold level for participating in the experiment; more details are in section 3 (on the participants) and the following section (on piloting and the data-collection instruments).

⁹¹ It should be mentioned that this 0% was only true if the five-score amount of residual uncertainty was correct.

4.3.6 Scoring: Procedure and Method

This section describes briefly the procedures and methods adopted in marking/scoring the participants' performance on both tasks: the GJ task and the translation task.

4.3.6.1 The GJ Task: The Adopted Marking Method

It is important, when considering the marking system used on the GJ test, to illustrate the fact that although the task required the participants to judge the test sentences on a four-point rating scale (*clearly correct*, *clearly incorrect*, *possibly incorrect*, and *I don't know*), their responses to each sentence were evaluated according to the following marking formula that covers all of their possible reactions to the sentence:⁹²

1. clearly correct (CC)
2. clearly incorrect, and the right correction was provided (CIT)
3. clearly incorrect, but a wrong correction was provided (CIF)
4. clearly incorrect, but no correction was provided (CIN)
5. possibly incorrect, and the right correction was provided (PIT)
6. possibly incorrect, but a wrong correction was provided (PIF)
7. possibly incorrect, but no correction was provided (PIN)
8. I don't know (DN)

⁹² Although the words *response* and *reaction* are often used interchangeably in the linguistics field, they will be used differently in this chapter. The word *response* will be used here only when referring to the four predetermined response options given to the participants in the GJ task. The word *reaction* will be used to refer to possible various ways the participants react to the given responses – the nine possible reactions presented above.

9. Missing response (NA)⁹³

This marking method, though it has not been used in published work, was implemented for the following reasons:

1. It allows the researcher to deal with the different reactions to the sentence that indirectly lead to the same concept/meaning, at least from the research goal's perspective. In other words, it enables the researcher, based on the correction provided, to regroup certain reactions that relatively share the same meanings, as follows:
 - i. Any ungrammatical sentence rejected based on a reason that was not related to its grammatical incorrectness, whether it was judged to be either clearly incorrect or possibly incorrect, will be considered as if it had been judged clearly correct.
 - ii. Any ungrammatical sentence judged to be possibly incorrect that was successfully rejected based on a reason related to its grammatical incorrectness will be counted as if it had been judged clearly incorrect.

Note that because of the criterion in (ii), the *possibly incorrect* option will not be part of the analysis as a distinct category.

2. It gives the researcher the opportunity to handle some of the methodological problems commonly associated with participants' responses to judgements such as the missing responses, the incomplete responses as expected/required, and the don't know responses that can cause either significant data loss or biasing of the results, depending on the method of analysis adopted by the researcher to deal with such pitfalls. These problems were addressed as follows:
 - i. Handling the missing reactions

Because it is not possible for the researcher to know whether a missing response is meaningful or not (e.g., whether the respondent left it out by mistake or omitted it intentionally because he or she did not want to

⁹³ I will consider in this study such missing values as a possible reaction to sentences. This is because respondents sometimes skip some questions intentionally, not by mistake, for various reasons (cf. Low, 1999; Dörnyei, 2003).

answer it or had no idea about the answer), any missing responses will be excluded/removed from the statistical analysis.⁹⁴

ii. Handling the don't know responses

Given that such *don't know* responses can provide no information about the different/similar routes/stages through which learners pass in developing L2 grammatical knowledge, all such *don't know* responses to ungrammatical experimental sentences will be excluded/removed from the statistical analysis.

iii. Handling the incomplete responses as instructed

Because it is extremely difficult to know whether a sentence judged as either clearly incorrect or possibly incorrect was rejected based on reasons related to its grammatical incorrectness without providing a correction to that sentence, and since the purpose of the present study is to compare and describe some characteristics of the different interlanguages of the same language – the Englishes of Arabic learners, French learners, and Finnish learners – all such sentences will be excluded/removed from the statistical analysis if the perceived or possible error in that sentence was not underlined or corrected.⁹⁵

3. It allows the researcher to exclude the participants who did not perform the task as expected in a very methodical way. The method used to exclude such

⁹⁴ In a few cases where some of the participants failed to judge the last items because they probably ran out of time, such unjudged sentences were removed from the statistical analysis as well.

⁹⁵ Since it is commonly noted that there is a relationship between number of errors made and proficiency level and between type of errors and L2 learners' mother tongue, it was sometimes possible for the researcher to figure out whether the uncorrected sentence had been rejected for the reason that made it ungrammatical, especially with advanced learners, who are expected to converge on native-like usage at the late stages of L2A, and especially when comparing uncorrected sentences with other similar rejected but corrected sentences. Since this procedure could produce inconclusive results affected by the researcher's own opinion, however, the decision made was not to include such sentences in the analyses.

participants was based on meaningful information about the links (a) between the possible lack of seriousness in completing the task and providing no correction to some of the rejected sentences or providing no responses at all to a number of sentences and (b) between lack of sufficient levels of proficiency in English to perform the task and providing too many don't know responses.⁹⁶ As a direct consequence of such links that could invalidate the task, the criterion formulated to exclude such participants from the analysis is as follows:

Any participant who had 20% (8 sentences) or more of his or her reactions to the experimental sentences (38 sentences) excluded from the analysis based on the criteria stated in (2) above will not be included in the study.⁹⁷

This adopted marking method requires, after marking the participants' reactions individually,⁹⁸ converting the performance on each sentence into a numerical score. Therefore, one point was given for each reaction a participant made, regardless of its correctness. After that, these points for each of the nine possible reactions were calculated for every participant using Microsoft's Excel 2010 programme. Then, to

⁹⁶ If for any reason the proficiency test failed to assign the participants to their correct levels, this procedure is likely to solve the problem. More details about the proficiency test and the threshold level required for participating in the study were given in section 4.4.

⁹⁷ It was necessary to implement this percentage-based exclusion plan because "it is quite common to have a few missing values in every questionnaire" (Dörnyei, 2003, p. 106). Otherwise I would end up losing a lot of data if any participant who did not complete all the task items as required were to be excluded from the analysis. After all, I think that the minimum remaining 30 sentences (80 per cent of the total number of the experimental sentences) was enough to reasonably answer the research questions of the present study. White (1985), for example, used 30 sentences in her experiment, and it has been one of the most important studies that have been conducted to investigate L2 acquisition of pro-drop parameter. It has been cited by 375 authors so far (see http://scholar.google.ca/citations?view_op=view_citation&hl=en&user=lgaiOYIAAAAJ&citation_for_view=lgaiOYIAAAAJ:UeHWp8X0CEIC).

⁹⁸ Not according to the four predetermined response options given to the participants but according to scoring the possible reactions discussed above.

calculate the percentage of the excluded responses for each participant in order to identify which participant(s) should be excluded from the statistical analysis according to the 20% excluding criterion discussed above, the participants' scores of the four excluded reactions (1. Clearly incorrect – no correction was provided, 2. Possibly incorrect – no correction was provided, 3. I don't know, and 4. Missing response) of only the experimental sentences were added together.⁹⁹ The following tables illustrate these marking, responses classification, and calculation processes:

Table 4-7. Methods used to mark the GJ test

Participant No.	Reactions to the 70 Test Sentences								
	S2	...	S20	S21	...	S50	S51	...	S70
FN P18	CIN	...	CIT	CIT	...	NA	NA	...	CIT
FN P49	CC	...	CIT	CIT	...	CC	CIT	...	CIF
FR P119	CIT	...	PIT	PIT	...	CIT	PIT	...	PIT
FR P141	CC	...	DN	PIN	...	PIN	DN	...	DN
AR P258	NA	...	PIF	PIT	...	PIT	CC	...	DN
AR P273	CC	...	CC	CC	...	PIN	CIN	...	DN

Key:

CC: clearly correct

CIT: clearly incorrect – the right correction was provided

CIF: clearly incorrect – a wrong correction was provided

CIN: clearly incorrect – no correction was provided

PIT: possibly incorrect – the right correction was provided

PIF: possibly incorrect – a wrong correction was provided

PIN: possibly incorrect – no correction was provided

DN: I don't know

NA: missing response

Table 4-8. Classifying the possible responses to the judged sentences: included responses vs. excluded responses

Classification		Possible Responses to the Sentences								
		CC	CIT	PIT	CIF	PIF	CIN	PIN	DN	NA
Included Responses	Acceptance	✓			✓	✓				
	Rejection		✓	✓						
Excluded Responses							✓	✓	✓	✓

⁹⁹ This is because data analysis of the GJ task will concentrate mainly on the participants' acceptance or rejection of only the ungrammatical sentence (for more details, see Chapter 5).

Table 4-9. Calculation processes used to exclude reactions and participants in the GJ task

Participants	Excluded Responses				Exclusion
	Null Subject	Missing Subject–Verb Agreement	Total	Percentage	
FN P18	6	2	8	20 %	Included
FN P49	3	2	5	13 %	Included
FR P119	0	0	0	00 %	Included
FR P141	13	7	20	53 %	Excluded
AR P258	2	0	2	05 %	Included
AR P273	17	7	24	63 %	Excluded

Tables 4.7, 4.8, and 4.9 above summarise, respectively, the procedure used to mark the GJ task, the calculation processes used to exclude responses, and the results used to exclude participants. For example, based on the 20% excluding criterion, only FR-P141 and AR-P273 would be excluded from the statistical analysis, as the total number of their excluded reactions exceeded eight reactions – more than 20% of the total number of the expected reactions of the experimental sentences.

As for the remaining five possible reactions (CC, CIT, CIF, PIT, and PIF) to the ungrammatical sentences that would be included in the statistical analysis, they were regrouped into only two categories (*acceptable English sentence* and *unacceptable English sentence*) based on the correctness/incorrectness of the correction provided to represent both the accepted and the rejected sentences. Table 4.10 explains the criteria used for regrouping them:

Table 4-10. Method used in regrouping the included reactions: accepted vs. rejected English sentences

Participants	Included Responses to the Experimental Sentences with Null Subjects				
	Total	Acceptable English Sentence (CC, CIF, PIF)		Unacceptable English Sentence (CIT, PIT)	
		Subtotal	%	Subtotal	%
FN P18	22	4	18	18	82
FN P49	25	9	36	26	64
FR P119	28	0	0	28	100
FR P141	15	7	47	8	53
AR P258	26	19	73	7	27
AR P273	11	11	100	0	0

This table shows that all the sentences rejected based on any reasons that were not related to their grammatical incorrectness will be considered as if they had been judged a clearly correct sentence.¹⁰⁰ This is evident from the fact that three responses (CC, CIF, PIF), as they have the same meaning, are added together to form the category *acceptable English sentence*. On the other hand, this table shows that all the reactions to the ungrammatical sentences judged *clearly incorrect* or *possibly incorrect* for which right correction was also provided would be added together to form the category *unacceptable English sentence*.

One issue that should be stated at this point is that all 70 test sentences have been classified into their subgroups, illustrated in subsection 4.3.2.1.2, prior to the marking and data-entry phases and using the same Excel programme. This classification process was done to further prepare the data sets for statistical analysis (see section 4.3.7).

In sum, this novel method of scoring seems to offer two major advantages.¹⁰¹ On one hand, by solving some of the methodological problems that can invalidate or at least bias the results of the GJ task, it ensures that for every student, firm reliable conclusions can be drawn that reflect his or her individual knowledge in detecting the

¹⁰⁰ One of the examiners wondered if considering any sentence that was rejected based on a reason that was not related to its grammaticality as if it had been judged as 'clearly correct' was the best idea. He drew my attention to the possibility that when a learner correctly identified a test sentence as ill-formed based on a reason that was not related to its grammaticality, we do not know for sure why he or she took that decision even though an incorrect correction was provided. What if a participant is capable of identifying sentences containing an error, but unable to correct them? Surely this would imply some knowledge of the target item. To deal with this matter, he suggested conducting an analysis to investigate whether individuals who failed to correct sentences were nonetheless able to identify incorrect sentences. However, it is hard to look at individual results due to the large number of participants involved in this study; therefore, only the group results were presented, examined, compared, and discussed in Appendix 13 on page 286-87.

¹⁰¹ Other marking methods used by researchers have been illustrated in subsection 4.3.2.1.2, when describing the various rate scales used by them and the problems associated with them.

ungrammaticality of null arguments in English in the contexts being investigated. On the other hand, by making methodical changes in the data set through excluding and regrouping certain reactions, it appropriately prepares the data for statistical analysis by making it easier to handle, yet more reliable. How this refined data was statistically analysed is the focus of section 4.3.7.

4.3.6.2 The Translation Task: The Adopted Marking Method

The method adopted to mark the translation task was similar, to some extent, to that already described in the previous section for the GJ task. They do differ from each other in certain ways, however. Given the nature of the translation task, which is designed to test the production of grammaticality, not the intuitive knowledge of grammaticality and ungrammaticality the way the GJ task did, the participant's translation to each sentence was marked on the basis of the following two marking criteria:

1. Whether he or she produced the relevant syntactic structure for the analysis or not – namely, whether the construction produced contains an embedded adverbial or complement clause or instead contains a clause or structure irrelevant to the analysis. Note that native speakers in such sentential contexts usually use a complex sentence with either an embedded adverbial or complement clause¹⁰²
2. Whether he or she managed to supply the pronominal functional category and/or the agreement morphology under investigation in this study – namely, whether he or she provided the required overt embedded subject pronoun and the subject–verb agreement inflection

¹⁰² It was noticed in the pilot studies that native speakers on rare occasion used the gerund form with some of the embedded adverbial clauses in the GJ task when correcting the ungrammatical sentences. Such translated construction will be considered as correct target translation. This is because using the gerund form in this particular position indicates that the learner has realised that English cannot have a null *pro* subject in this syntactic position.

This marking method was carried out, again using Microsoft Excel 2010, so that each translated sentence was evaluated according to the following five marking options that cover all the learners' possible translation performances:

1. Target-like structure – correct suppliance of the investigated form (CS)
2. Target-like structure – incorrect suppliance of the investigated form (IS)
3. Non-target-like structure/unexpected translation (DS)
4. Unanalysable translated sentence (US)
5. Missing translation (NA).

The rationale behind implementing this marking method was again to ensure that the performance for every participant is accurate enough to meet the purpose the test was designed for – to reflect the participant's IL. To meet this purpose, the results must be reliable. Ideally, the principles of this marking method enhance the reliability of the data obtained by making it possible for the researcher to properly deal with the data that can bias or invalidate the findings of the study. For this reason, all translational performance that neither can add information about the learner's IL (i.e., *unanalysable-translated sentence* [US] or *missing translation* [NA]) nor is structurally relevant to the study (i.e., *non-target-like structure* [DS]) will be removed from the statistical analysis. The following table illustrates which of these possible translational performances are including in the statistical analysis and which are not.

Table 4-11. Possible translational performances: included performances vs. excluded performances

Classification	Participants' Possible Translational Performances				
	CS	IS	DS	US	NA
Included Performances	✓	✓			
Excluded Performances			✓	✓	✓

Key:

CS: Target-like structure – correct suppliance of the investigated form

IS: Target-like structure – incorrect suppliance of the investigated form

DS: Non-target-like structure/unexpected translation

US: Unanalysable translated sentence

NA: Missing translation

Furthermore, this marking method allows the researcher to exclude participants who did not complete the task as expected in a convincing, systematic way, following

the principles of the 20% excluding criterion formulated in the GJ task. Consequently, any participant whose excluded performance (US, NA, DS) exceeded 20% of his or her translational performance (4+ out of 15 sentences) was removed from the analysis.

These answer codes (CS, IS, US, NA, DS) used in the marking process need to be transformed into numerical scores for the purpose of statistics. This process of transforming answers into numbers was done by following the same process performed to deal with the participant's reactions to the grammatical and ungrammatical sentences in the other task, where initially the participant was given one point for each evaluated translational performance, regardless of its structural relevance or correctness. After that, these points for each type of the five possible translational performances were calculated for every participant using Microsoft's Excel 2010 programme. Then, to discover whether a participant should be excluded from the analysis according to the 20% excluding criterion, the participant's scores for the three excluded translational performances (US, NA, DS) were added together. The following tables offer an overview of these marking, calculation, and exclusion processes:

Table 4-12. Method used to mark the translation test

Participant No.	Translational Performances to the Test Sentences						
	S1	S2	S3	...	S13	S14	S15
FN P18	CS	CS	IS	...	NA	CS	CS
FN P36	CS	CS	CS	...	NA	NA	NA
FR P119	CS	CS	CS	...	CS	CS	CS
FR P175	CS	CS	CS	...	NA	NA	NA
AR P258	CS	CS	CS	...	CS	CS	CS
AR P283	IS	CS	CS	...	CS	CS	NA

Key:

CS: Target-like structure – correct suppliance of the investigated form

IS: Target-like structure – incorrect suppliance of the investigated form

DS: Non-target-like structure/unexpected translation

US: Unanalysable translated sentence

NA: Missing translation

Table 4-13. Calculation procedures used to exclude performances and participants in the translation task

Participants	Number and Percentage of the Excluded Translational Performances		Exclusion
	Null Subject (Non-TS, UTS, MT)	Percentage	
FN P18	1	07 %	Included
FN P36	6	40 %	Excluded
FR P119	0	00 %	Included
FR P175	4	27 %	Excluded
AR P258	0	00 %	Included
AR P283	1	07 %	Included

These tables summarise the different processes involved in marking the translation task. Table 4.12 shows how the different translational performances were initially marked, whereas Table 4.13 illustrates the implemented participant-excluding criterion (i.e., participants FN-P36 and FR-P175 were removed from the analysis as their excluded individual performance [US, NA, DS] exceeded 20% of their total translational performances). Accordingly, it implicitly exhibits that only the CS and IS translational performances will be part of the analysis. Hence, it could be argued that this coding and marking procedure not only prepares the data for the next descriptive and inferential statistical step but also increases the reliability of the data, allowing conclusive conclusions and generalisations to be drawn from this study.

4.3.7 Data Analytical and Statistical Framework

This section will illustrate in two ways how the learners' data was analysed: The first one has to do with the methodological approach – the analytical procedure – adopted in examining and comparing the data of the participants. The second has to do with the statistical software tool used in analysing the quantitative data and the various statistical techniques utilised to make the necessary reports, comparisons, and contrasts.

Given the complex nature of the interlanguage research questions stated in 4.2.2, it was important for the researcher to adopt a reliable and valid grammatical-error comparison approach to examine learners' L2 use in order to generate

convincing and generalizable answers to the study questions, such as the renewed version of the traditional contrastive analysis (CA) approach referred to by Granger (1996, 1998) as the *contrastive interlanguage analysis* (CIA).¹⁰³ Instead of comparing the native and the target languages of learners, the CIA compares and contrasts “what non-native and native speakers of a language do in a comparable situation” (Granger, 1998, p. 12, after Pery-Woodley, 1990, p. 143). This comparative model involves basically two major types of comparison: a comparison between native speakers and L2 learners (native language vs. interlanguage) to “uncover the patterns of use distinguishing learner data from native data” (Granger, 2003, p. 541), and a comparison between L2 learners of the same language with different L1 backgrounds (ILs vs. ILs) “to establish whether the differences uncovered are developmental or transfer related” (Granger, 2003, p. 541). The former type may also involve comparing L1 child data with adult L2 learner data to uncover the similar and/or different patterns of acquisition between these different learner groups. The latter may also involve comparing data of L2 learners of the same language who share the same L1 background but are at different stages of L2 development in order to identify the characteristics of different IL stages.

It can be claimed, on the basis of such different systematic analyses of L2 learners’ data, that this method of comparison is very reliable when conducting interlanguage research. It has not only deepened our understanding of the nature of interlanguage by answering many fundamental questions in SLA research, but it also has opened up various unexpected avenues of enquiry into the field of study. Granger (2009) points out that such “L1–L2 comparisons are extremely powerful heuristic techniques which help bring to light features of learner language which have not been focused on before, and which, once uncovered, can be analysed from a strictly L2 perspective” (p. 18).

To make such necessary complex analytical comparisons, a relatively new but appropriately flexible statistical tool was applied in the data analysis – the R Statistical

¹⁰³ For a detailed historical review of the origin of the CIA, which was originally referred to as a “new type of CA” by Selinker (1989, p. 14), see Gilquin (2001).

Programme.¹⁰⁴ Like SPSS, R can execute sophisticated statistical analyses by providing a wide variety of statistical procedures. Not only is it a completely free and easy-to-learn open-source program, but one of R's strengths over SPSS and many other famous statistical packages is its graphing capabilities; it provides hundreds of effective ways to present the data in very beautiful and reader-friendly data depictions. Over the past few years, it has become the statistical programme of choice for many researchers in a variety of fields including linguistics.¹⁰⁵

Prior to running the statistical analyses, however, the refined Excel data sets obtained from the participants' performance on both tasks discussed in 4.3.6 were imported into R. Then, for the purpose of descriptive statistics, each subgroup's results were reported including information about the mean, minimum, maximum, standard deviation, and percentage of acceptance or omission for each variable under investigation. These quantitative values were calculated out of the total number of the included accepted ungrammatical items in the GJ task and of the included omitted target form in the translation task, except for the percentages of acceptance and omission, which are calculated by dividing each sum of acceptances or omissions by the total number of the included responses.¹⁰⁶ After that, to pave the way for running the appropriate statistical inferential procedures, the sample's normality of distribution was checked using the Shapiro test. Not all of the subgroups' results in the two tasks were normally distributed ($p > .05$) for all variables.¹⁰⁷

¹⁰⁴ R was initially developed by Robert Gentleman and Ross Ihaka at Auckland University. To obtain more comprehensive information about R, visit its official website: <http://www.r-project.org/>

¹⁰⁵ For more information about how to use R to process linguistic or psycholinguistic data, refer to Baayen (2008) and Gries (2009).

¹⁰⁶ Refer back to subsections 4.3.6.1 and 4.3.6.2 to see how the included responses and omissions were calculated in both of the elicitation tasks.

¹⁰⁷ Because of the large number of the participants' subgroups and the large number of investigated variables in both tasks, and since normality tests are applied only to select the appropriate parametric or nonparametric statistical techniques to be used with the data, the results of such tests will not be presented or discussed in this chapter. A list of

Therefore, both parametric and nonparametric procedures were used. Accordingly, seven types of statistical tests were applied to test for significant differences between and within the subgroups of learners and to examine the strength and direction of the relationship between variables. If the data are normally distributed, (a) paired samples *t*-tests are used when two variables are compared within the same subgroup, (b) independent samples *t*-tests are used when the results of two subgroups for the same variable are compared against one another, and (c) analysis of variance (ANOVA) followed by a post hoc test (Tukey HSD) are used when the results of more than two subgroups for the same variable are compared against one another. If the data is not normally distributed, however, nonparametric tests are used instead of the parametric ones; for example, (d) the Wilcoxon test, which in R is called *Wilcox.test*, is the nonparametric alternative test to both the paired samples *t*-test and the independent samples *t*-test (but used in two different ways), and (e) the Kruskal-Wallis test, in R called *Kruskal.test*, is the nonparametric alternative test to the ANOVA.¹⁰⁸ As for the correlation analysis, (f) the Pearson product–moment correlation is run with both normally and not normally distributed data to compute the correlation coefficient between two variables; it is abbreviated as *r*.¹⁰⁹ As for the interaction analysis, (g) the generalised linear mixed effects model (GLMM) was used to investigate whether the performance on the variables under investigation differ depending on the learners' L1 and proficiency levels. The alpha level was set at $p < 0.05$ for all of these tests; in other words, a result is considered significant if $p < 0.05$.

the results of a total of 126 normality tests applied for all variables and samples is available in appendix 8.

¹⁰⁸ Unlike the R Statistical Programme, SPSS uses two different nonparametric tests instead of the two types of *t*-tests: the Mann-Whitney U test is the alternative to the independent samples *t*-test, and the Wilcoxon signed-rank test is the alternative to the paired samples *t*-test. Refer to Dörnyei (2007) for the SPSS and to Baayen (2008), Gries (2009), and Race (2012) for R.

¹⁰⁹ The Spearman's rank order correlation is designed to compute *r* for those data that do not satisfy the distribution normality. Nevertheless, it was not used here because its result is less powerful than Pearson's, which can also be used for the data not normally distributed (see Dörnyei, 2007; Race, 2012).

The next chapter illustrates precisely how these descriptive and inferential statistical procedures are put into practice.

Chapter 5. Results and Discussion

5.1 Introduction

The primary goals of this chapter are (1) to present the empirical data that emerged from the data elicitation tools, (2) to interpret and discuss these results in light of the theoretical assumptions outlined in the previous chapters, and (3) to connect the findings to those of previous studies in this field.¹¹⁰ However, because the research questions and hypotheses formulated in the previous methodology chapter revolve around the issue of transfer of the L1 value of the null subject parameter, and in order for transfer to be unambiguously established and understood, only the group results from non-target-like performance relating to subject pronouns as well as missing verbal agreement are presented, examined, compared, and discussed here. That is to say, the data analysis and discussion of the results will concentrate on the participants' non-well-formed translated items and on their acceptance of sentences that are ungrammatical with respect to the syntactic properties being investigated.¹¹¹ The only exception is when the participants' strength of preference for overt pronouns over null forms needs to be measured; in Subsection 5.2.3.2, I will look at the participants' rejection of these ungrammatical items.

Note, as briefly mentioned above, I will only analyse, report, and discuss the group results in this chapter; the individual results will not be analysed here, except in one context when discussing the notion of parameter resetting in Subsection

¹¹⁰ Because the analysis and the discussion are closely connected, a decision has been made to have them in one chapter against the norm, which is two separate chapters. Such a procedure helps the author to discuss the results in a cohesive way without making many references throughout this chapter to the findings.

¹¹¹ It's clear that the other percentage representing their performance is what they got right/target-like.

5.3.2.2, to illustrate the fact that the group results cannot be applied to all the individuals in each group. However, due to the large number of participants and space limitations, only the descriptive results from the advanced Finnish-speaking participants are presented at the individual level to demonstrate the problem of generalisation in L2 development.

This chapter comprises four sections to address the research questions listed in 4.2.2. Section 5.2 investigates the results of the analysis of the data and seeks to answer the first and second research questions, namely if null subject transfers in L2A equally affect grammatical intuitions and oral production and if parameters can be reset in L2A. Section 5.3 investigates the data to answer the research question regarding whether the results of the subgroups diverge similarly or differently across the different syntactic formations; it also reconsiders the second research question, namely if the L1 value of the null subject parameter can be reset in L2A. Section 5.4 tries to answer the third research question; it seeks to investigate the mechanisms by which the presence of a null subject is licensed in the learner's IL grammars.

5.2 Results and discussion by task

This section, which presents and discusses the results as they emerge from the data collection tools to see if the learners, as subgroups, would transfer their L1 setting of the null subject parameter, consists of three subsections. The first subsection deals with the results of the GJ task; it presents descriptive and statistical analyses of the results of the GJ task. Namely, the quantitative differences between each L1 group (French, Finnish, and Arabic) and its subgroups (lower-intermediate, upper-intermediate, and advanced) will be examined, compared, and contrasted using a series of inferential tests against the native-speakers control group as well as against one another, following the principles of Contrastive Interlanguage Analysis presented in subsection 4.3.7. Therefore, it will include three sets of inter-subgroup comparisons: (i) between the L2 learners and the control group to find out how the ILs diverge from the grammar of English native speakers, (ii) among L2 learners from the same L1 background at different developmental stages to find out how ILs change

over time and whether the ILs at the final stage of development converge on L2 grammar, and (iii) among L2 learners from different L1 backgrounds who are at the same proficiency level to distinguish whether the differences observed among the different L1s subgroups are developmental or transfer related. The second subsection deals with the results of the translation task in the same way the results of the GJ task were analysed and interpreted in accordance with the principles of Contrastive Interlanguage Analysis. In the third subsection, I examine how the results diverge within each subgroup across the two tasks. To make such intra-subgroup comparisons, I compare and contrast each subgroup's performance on the GJ task statistically with their translation task to examine if the learners' performances vary from task to task. The findings of this subsection will lead the discussion to issues of parameter resetting and ultimate attainment.

Note that the notion of intergroup comparison is different from the notion of intragroup comparison. The former is conducted to describe the variation/difference or even similarity among L2 (sub)groups of learners, whereas the latter is done to investigate the degree of variability or inconsistency in data from the same group of learners but in different performance contexts or occasions.

5.2.1 The grammaticality judgment task

5.2.1.1 The English ILs of the French, Finnish, and Arabic speakers: Basic descriptive and inferential statistics

i. The English IL: French speakers

The descriptive statistics of the GJ task data for all three subgroups of French-speaking English L2 learners are presented in Table 5.1. It reveals that the French lower-intermediate learners, as a subgroup, incorrectly accept far more sentences with missing embedded subjects than the upper-intermediate and advanced subgroups. Notably, the lower-intermediate learners' mean acceptance score was quite high (12.44, StdDev 3.57) in comparison with the upper-intermediate learners' (3.00, StdDev 2.03) and the advanced learners (0.52, StdDev 1.12).

Table 5-1. Descriptive statistics for French-speaking learners' acceptance of ungrammatical sentences with null embedded subjects

Proficiency level	Included responses	No. of Acceptance	Max	Min	Mean	StdDev	Acceptance percentage
LI (N= 09)	227	112	18	8	12.44	3.57	49.3
UI (N= 31)	802	93	8	0	3.00	2.03	11.6
ADV(N= 27)	748	14	5	0	0.52	1.12	1.8

N= number of participants

To see if the learners' proficiency levels affect their performance with respect to accepting sentences with null embedded subjects, the subgroups' results were submitted to the Kruskal Wallis Test. The test result shows that there is a statistical significant difference among these proficiency subgroups (P-value = 0.000<0.05).

However, when the results in Table 5.1 above are statistically compared with the results of the English native controls discussed in Subsection 4.3.3.2 and descriptively presented below in Table 5.2 via the Wilcoxon tests, the inferential results show that the French speaking learners differed significantly from native English speakers with respect to acceptance of null referential subjects at lower-intermediate and upper-intermediate levels only. The analytical results as compared with native controls were: P-value = 0.001<0.05, P-value = 0.000<0.05, and P-value = 0.148>0.05, respectively, for the lower-intermediate subgroup, upper-intermediate subgroup, and advanced subgroup.

Table 5-2. Descriptive statistics for the English native-speakers control group: acceptance of ungrammatical sentences with null embedded subjects

Number of participants	Included responses	No. of acceptance	Max	Min	Mean	StdDev	Acceptance percentage
7	194*	0	0	0	0.00	0.00	0.0

* The included responses = number of the relevant experimental sentences in the task x number of the participants – number of the excluded items (28 x 7 – 2 = 194).

The results presented so far suggest that the L1 parameter setting does not transfer at the initial stages of L2 development; the lower-intermediate French-speaking participants were found to accept null subjects in English despite the fact that neither their L1 nor their L2 allows null subjects. However, the performance of the advanced-level French participants which was native-like may suggest that the L1 setting is transferred at the late stages of L2A. These findings will be discussed later

in detail in subsections 5.2.1.2 and 5.2.1.3, where all of the subgroups' results from the GJ task are compared and discussed.

ii. The English ILs: Finnish speakers and Arabic speakers

The overall performance of the L1 Finnish learners and L1 Arabic learners exhibits relatively similar patterns of development to their L1 French-speaking counterparts, but with varying degrees of accuracy in relation to the acquisition of overt subject pronouns in English. The lower-intermediate subgroups accept noticeably more ungrammatical sentences with null embedded subjects than do the upper-intermediate and advanced subgroups. As Table 5.3 shows below, the means of acceptance of the L1 Finnish participants, on the one hand, are 11.67 (StdDev 2.66), 3.12 (StdDev 2.76), and 0.64 (StdDev 0.86) for the lower-intermediate learners, upper-intermediate learners, and advanced learners, respectively, whereas their Arabic counterparts accept similar ungrammatical items at the mean of 16.81 (StdDev 3.66), 6.86 (StdDev 3.61), and 1.27 (StdDev 1.42) for the lower-intermediate learners, upper-intermediate learners, and advanced learners as subgroups, respectively.

Table 5-3. Descriptive statistics for L1 Finnish- and Arabic-speaking learners' acceptance of ungrammatical sentences with null embedded subjects

Proficiency level	Included responses	No. of acceptance	Max	Min	Mean	StdDev	Acceptance percentage
Finnish LI (N= 06)	152	70	15	9	11.67	2.66	46.0
Finnish UI (N= 17)	450	53	9	0	3.12	2.76	11.7
Finnish ADV (N= 53)	1472	34	3	0	0.64	0.86	2.3
Arabic LI (N= 16)	397	269	24	12	16.81	3.66	67.7
Arabic UI (N= 14)	361	96	13	1	6.86	3.61	26.5
Arabic ADV (N= 11)	299	14	4	0	1.27	1.42	4.6

Similar to the L1 French speakers' performance on null subject pronouns, the differences in performance between the L1 Finnish lower-intermediate-level, upper-intermediate-level, and the advanced-level learners and between the Arabic lower-intermediate-level, upper-intermediate-level, and the advanced-level learners are all

statistically significant by the Kruskal Wallis tests, which is $P\text{-value} = 0.000 < 0.05$ for the Finnish and Arab learners.

However, unlike the French-speaking learners, specifically in the case of the advanced-level subgroup, when the Finnish and Arabic participants' subgroups' results presented in Table 5.3 above are statistically compared with the results of the controls in Table 5.2, all the results obtained via the Wilcoxon tests indicate that the L1 Finnish and L1 Arabic participants of all proficiency levels differed significantly from native English speakers. As subgroups, these L2 learners continue to perform non-native-like with respect to acceptance of null referential subject pronouns in English despite the fact that their performance improves with increased proficiency in the L2. Consider the inferential results illustrated in Table 5.4.

Table 5-4. Inter-subgroup inferential comparisons between the L2 learners (L1 Finnish and L1 Arabic) and the native-English controls

Inter-subgroup comparisons	Proficiency level	Inferential test	P-value ¹¹²
L2 Finnish participants vs. native English controls	LI	Wilcoxon	0.034<0.05
	UI	Wilcoxon	0.001<0.05
	ADV	Wilcoxon	0.001<0.05
L2 Arabic participants vs. native English controls	LI	Wilcoxon	0.027<0.05
	UI	Wilcoxon	0.000<0.05
	ADV	Wilcoxon	0.000<0.05

The results presented in this section suggest that the L1 setting transfers at the initial stages of L2 development; both the L1 Arabic- and Finnish-speaking L2

¹¹²One of the examiners suggested that use of star ratings for p-values would make tables easier to interpret, e.g. * < 0.05, ** < 0.01, *** < 0.001. I have not followed this recommendation because I think giving/writing the actual results gives an accurate reading of the P-values. Moreover, the suggested method (use of star ratings) causes confusion about what level of confidence is set in the study to consider whether a result is statistically significant or not; in other words, this method could confuse the reader about whether the alpha level was set at $p < 0.05$, $p < 0.01$, or $p < 0.001$. In the present study the alpha level was set at $p < 0.05$ for all of the tests applied (see section 4.3.7). Note, in order to make the tables easier to interpret, I highlighted in grey all the results (P-values) that show statistical significance (see the tables in Chapter 5).

learners of English were found to accept null subjects in English. However, this cannot be taken for certain as the lower-intermediate French-speaking participants were found to accept null subjects in English despite the fact that neither their L1 nor their L2 allow null subjects. In the following section, I will compare the developmental paths – based on the cross-sectional data - of these L2 learners of English of typologically different L1s before any conclusions can be drawn from the results presented so far with regard to whether or not the L1 setting transfers at the initial stages of L2 development.

5.2.1.2 Comparisons across all language subgroups: ILs vs. ILs of the same proficiency level

Although the comparisons conducted above among the different subgroups of proficiency levels (lower-intermediate, upper-intermediate and advanced) within each L1 group of learners show that the higher the proficiency level, the better at rejecting ungrammatical sentences with null subjects, we cannot assume that starting off by accepting ungrammatical sentences with a null subject is a developmental phenomenon affecting all L2 learners regardless of their L1 background until we “compare the developmental paths of L2 speakers of typologically different L1s. If there is divergent development, then this [may] constitute evidence for transfer” (Hawkins, 2001b, p. 354). In contrast, if a similar pattern of development is noticed among learners from different linguistic backgrounds, then this constitutes evidence for UG. For better comparative visualization, the results in Tables 1 and 3 are summarised below in Figure 5.1 in terms of means of ungrammatical acceptance.¹¹³

¹¹³ One of the examiners suggested that the use of line graphs may have been a better way to show group performance. I have not followed this recommendation because I think that such R boxplots manage to present the complex data in a comparatively simple and meaningful way. Given to the large number of sub-groups being compared (nine sub-groups in Figure 5.1 above), the use of such boxplots makes figures easier to read and interpret compared to the use of line graphs. This is true especially in cases when the performance of the nine sub-groups of participants included in the present study are

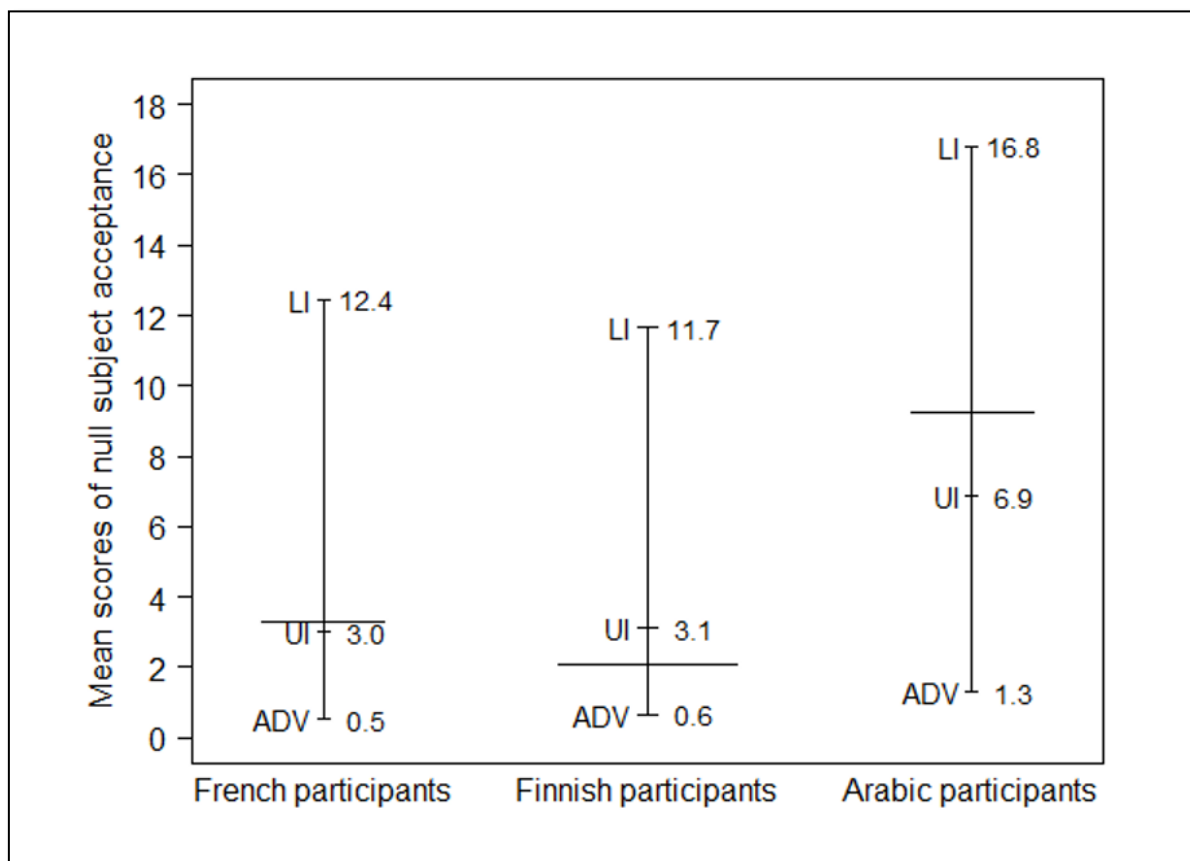


Figure 5-1. Summary and comparison of the subgroups' results in terms of means of null subject acceptance.

Therefore, inferential tests were conducted for comparisons between the L1 French, L1 Finnish, and L1 Arabic learners of the same proficiency level to determine whether the source of their performance was the product of their L1 backgrounds or UG. A one-way ANOVA revealed that there is a statistically significant difference with respect to accepting sentences involving null referential subjects between the lower-intermediate subgroups of the L1 French, Finnish, and Arabic learners ($P\text{-value} = 0.0032 < 0.05$). Similarly, a Kruskal Wallis test showed a significant difference between upper-intermediate subgroups for the same variable ($P\text{-value} = 0.0021 < 0.05$); however, there was no significant difference between the advanced

compared for two variables as in Figure 5.17 on page 202. Note, such boxplots are only used for better comparative visualisation; the groups' results are reported in tables including information about the mean, minimum, maximum, standard deviation, and percentage of acceptance or omission for each variable under investigation.

subgroups ($P\text{-value} = 0.1305 > 0.05$) – a finding which suggests that interlanguage (IL) restructuring occurs with time and increased proficiency.

It should be mentioned for the sake of illustration that there is a noticeable difference between the non-target performance of L1 Arabic learners and their French and Finnish counterparts. It is clear from Figure 5.1 above that L1 Arabic learners generally accepted more sentences with null embedded subjects than did their French and Finnish counterparts of the same proficiency level. It is, however, surprising that the performance of French learners as subgroups was comparable to that of lower-intermediate, UP, and advanced Finnish learners though their L1s are different in terms of parameter setting; they accepted about the same number of ungrammatical sentences at mean scores of 12.44, 3.00, and 0.52 for the French lower-intermediate, upper-intermediate, and advanced subgroups, respectively, and 11.67, 3.12, and 0.52 for lower-intermediate, upper-intermediate, and advanced Finnish learners, respectively. These observations are supported statistically by the results of the inferential tests performed as shown in Table 5.5; the results of the French vs. Finnish comparisons are shaded.

Table 5-5. The results of the statistical tests comparing the different subgroups of the participants on acceptance of null subjects

Proficiency level	Subgroups compared	Inferential test	P-value
ADV	Finnish–French–Arabic	Kruskal-Wallis	0.131 > 0.05
	Finnish–French	Wilcoxon	0.201 > 0.05
	Finnish–Arabic	Wilcoxon	0.185 > 0.05
	French–Arabic	Wilcoxon	0.067 > 0.05
UI	Finnish–French–Arabic	Kruskal-Wallis	0.002 < 0.05
	Finnish–French	Wilcoxon	0.785 > 0.05
	Finnish–Arabic	Wilcoxon	0.005 < 0.05
	French–Arabic	T-Test	0.002 < 0.05
LI	Finnish–French–Arabic	ANOVA	0.003 < 0.05
	Finnish–French	Tukey HSD	0.906 > 0.05
	Finnish–Arabic	Tukey HSD	0.012 < 0.05
	French–Arabic	Tukey HSD	0.014 < 0.05

Table 5.5 shows that the tests revealed that the differences between the L1 Arabic learners and L1 French learners and between the L1 Arabic learners and L1 Finnish learners are all significant at both lower-intermediate and upper-intermediate proficiency levels with respect to their performance on the variable under discussion. In contrast, the statistical tests showed that the differences between the L1 French

learners and the L1 Finnish learners of the same proficiency level were not in fact significant among all the proficiency subgroups. In other words, the data reveal that the Arabic-speaking learners accepted null subject sentences at a much higher rate than their French and Finnish counterparts, who unexpectedly accepted null subjects in a somewhat similar fashion.

Given how the three languages differ from each other in terms of the possibility of allowing null subjects in tensed clauses, the noticeable differences between the non-target-like performance of Arabic learners whose L1 allows subject-drop freely and their Finnish counterparts whose L1 allows subject-drop optionally in some contexts and excludes them in other contexts and between the non-target-like performance of Arabic learners and their French counterparts whose L1 does not allow subject-drop can be attributed to transfer from the L1. However, the L1-based knowledge influence cannot explain the similarities in performance between the L1 Finnish learners and the L1 French learners. Yet it can be argued that the Finnish participants accepted null subjects in a somewhat similar fashion as their French counterparts because Finnish allows null subjects in a restricted manner. This would lead to their preference for overt pronouns over null forms. Nevertheless, L1 transfer cannot explain why all of the lower-intermediate subgroups of learners, including the French, accept ungrammatical sentences with null embedded referential subjects despite the fact that their L1s are different in parameter setting (full non-pro-drop French, partial pro-drop Finnish, full pro-drop Arabic). This will be discussed in detail in the following section.

5.2.1.3 Discussion of grammaticality judgement tasks' results

Numerical and statistical results from the GJ task show evidence that all learners start off with a [+] pro-drop setting; namely, all the subgroups of learners regardless of their linguistic backgrounds accept ungrammatical sentences with null embedded referential subjects. They do so even at the latest stage of English learning, with the exception of the advanced-level French participants, who manage to converge on a native-like grammar. These findings, when connected to those of previous studies, are consistent with the common observations that emerge from the grammatical

intuition data (i.e., Tsimpli and Roussou, 1991; White, 1985, 1986)¹¹⁴ that native speakers of null subject languages who are L2 learners of non-null subject languages accept null subjects, as they do here - the Finnish and Arabic participants' intuitional results. On the other hand, they stand in contrast to White (1985, 1986), who argued that French-speaking learners of English (native speakers of non-null subject language who are L2 learners of a non-null subject language) do not show evidence of a [+prodrop] grammar.

This contrast is highly unexpected; in fact, White's result is the predicted one, in agreeing with the assumption that L1-based knowledge influences the IL (see the discussion about the Full Transfer/Full Access Hypothesis in Chapter 3). However, it has been observed before on GJ tests that French speakers accept null subjects in L2 development (e.g., see Liceras, 1989). Moreover, it has been observed that the ratio of null subject utterances in child's French can be high; statistically speaking, it is higher than what is reported in child English; it is almost comparable to that of child Italian (cf. Valian, 1990; Prévost, 2009, and the discussion in section 2.4 above). So, the preliminary question to be raised now is how the French-speaking learners accept null subjects. In order to answer this question in a meticulous, detailed way, first White's study (1985) that was presented in subsection 3.3 will further be critically reviewed, and then the results of the present study are submitted to further analysis to further explore the issue.

Apart from the methodological problems discussed in subsection 3.3 above that can make any results questionable, White analysed data in a way that makes one remain sceptical about her conclusion that L1 French speakers do not accept null subjects in L2 English. When White analysed the French results by levels of proficiency, she only gave the number and the percentage of acceptance for each individual sentence, as presented in Table 5.6 below.

Table 5-6. White's results (1985): French responses by level to individual sentences with missing subjects: number responding "correct." to ungrammatical sentences

¹¹⁴ With respect to White's studies, this is only true when the Spanish results are considered; see the critical comments on White below.

Level	Sentence Number					
	2	4	8	21	22	30
1–2 (N=5)	1 (20%)	2 (40%)	1 (20%)	1 (20%)	2 (40%)	3 (60%)
3 (N=8)	1 (12%)	4 (50%)	0	1 (12%)	0	1 (12%)
4–5 (N=6)	2 (33%)	1 (17%)	0	0	1 (17%)	0

N= number of French participants

These acceptance rates in Table 5.6 indicate that French learners are, to a certain extent, inclined to accept missing subjects in English. However, to get a more accurate picture, the percentage of null subjects acceptance for each level of proficiency (or *levels*, as White regrouped her participants, as indicated in point ii above) is calculated based on the information given in Table 5.6 above and illustrated in Table 5.7 below.

Table 5-7. White's results (1985): French responses by level to all sentences with missing subjects: percentage of acceptance

Level	Number of Participants	Number of Items	Number of Tokens*	Total Acceptance	Percentage of Acceptance**
1–2	5	6	30	10	33.3%
3	8	6	48	7	14.5%
4–5	6	6	36	4	11%

*Number of tokens = number of the participants x number of the test relevant item

**Percentage of acceptance = (total acceptance / number of tokens) * 100

It can be concluded based on these percentages of acceptance, especially the one of the level 1–2, that the French accept sentences with missing subjects. Actually, this conclusion is not only supported by the findings of the present study but also by Liceras's (1989) and Liceras, Díaz, and Maxwell's (1999) findings, which showed evidence that French learners of Spanish as an L2 not only accepted but also produced null subjects from early on. In fact, it is expected that, if null subjects are not part of French participants' competence, they would not accept pronoun omission in Spanish (see Liceras, 1989; Liceras, Díaz, and Maxwell, 1999) even though it is a null subject language, at least in the initial state of development.

These results raise the question of why French-speaking learners, especially at the lower-intermediate and upper-intermediate levels, accept null subject sentences despite the fact that French is arguably considered a non-null subject language. One possible answer is that the presence of null subjects in the French participants ILs is

a direct consequence of having an L1 that has subject clitics that form a morphological unit with the verb.¹¹⁵ To confirm this possibility, it would be better to compare the results of the French-speaking learners with intuitional results of another group of learners whose native language is a non-pro-drop language that does not have subject clitics, such as Swedish, learning a non-pro-drop language like English. However, I am not aware of any such study using a grammatical intuitional task that has been done at this point. Another plausible answer to account for the presence of null subjects in the French participants ILs could be that the French speakers did not transfer their L1 setting at the initial and intermediate stages of L2 development; rather, they began with a pro-drop grammar. Clearly, further research is required to investigate this issue in depth before a firm conclusion is drawn.

To summarise the grammatical intuitional results discussed so far, the data show that all the groups of learners—French, Finnish, and Arabic—accepted missing subjects in English. Despite the fact that gradual improvements in their abilities to recognize the ungrammaticality of such sentences came with increased proficiency in the L2, both the Arabic and the Finnish participants continued to perform in a non-target-like way; only the French participants managed to converge on a native-like grammar at the advanced-level. These results seem to disconfirm hypothesis H1 and its predictions (P1A, P1B and B1C), which were formulated based on the principles of the FT/FA hypothesis which sees L1 syntactic knowledge as the default source of L2 initial-state syntactic structures and the transitory intermediate states of syntactic knowledge. In fact, such results, which show that all the subgroups of learners accept ungrammatical sentences with null embedded referential subjects at the initial stages of L2 development despite the fact that their L1s are different in parameter setting (full non-pro-drop French, partial pro-drop Finnish, full pro-drop Arabic), are consistent with the Organic Grammar account of L2 development (Vainikka and Young-Scholten, 2011). This approach predicts that all L2 learners, regardless of their linguistic backgrounds, begin with a pro-drop grammar since their initial

¹¹⁵ For a detailed discussion about the different types of subject pronouns in French and their distributional properties, see Hawkins, Towell, and Lamy (2001); Prévost (2009); and Rowlett (2007).

grammar lacks functional projections; in other words, L2 learners begin with a bare VP projection in their IL grammar so that pro-drop parameter settings cannot be represented or transferred at the initial stages of L2A. However, the Organic Grammar approach, similar to the FT/FA hypothesis, predicts that at the final stage of development both the Finnish and Arabic learners will converge on the English L2 pattern [-pro-drop] if there is sufficient input. Therefore, Organic Grammar cannot explain why only the French participants, but not the Arabic and the Finnish ones, managed to converge on a native-like grammar at the advanced-level. However, the Modulated Structure Building approach (Hawkins, 2001a) may explain this, since it predicts that there is later transfer of information encoded in functional projections when they are posited; in other words, the L1 setting is transferred at later stages of L2A. The superiority of this account to all other approaches in explaining the results obtained from the GJ task is that it supports an initial state bare VP in L2A and at the same time allows for influence of the L1 during subsequent stages of development.

Before moving to the following section to discuss the data obtained by the translation task, the intuitional data are submitted to further analysis to further explore the nature of the null subject parameter in L2 acquisition and to make sure that this similarity in the performance of the French and Finnish learners was not due to an uncontrolled variable (i.e. multi-competence) or to any other reason that is not entirely clear so far. Thus, I further analysed the responses to the ungrammatical sentences based on the position of the potential referential antecedents of the embedded missing subjects—namely the antecedents within the sentence in the main clause (henceforth *local antecedent*) or the antecedents outside the sentence (henceforth *non-local antecedent*)—expected in broader discourse context (for more detail, see subsection: 4.3.2.1.2). The percentages of incorrect judgments of these items for the French and Finnish learner subgroups are represented in Figure 5.2.¹¹⁶ As Figure

¹¹⁶ Due to space limitations, and the fact that this type of analysis was conducted mainly to further test the prediction of hypothesis H1, before rejecting it based on the results presented in the previous section, detailed descriptive statistics for this analysis will not be presented here. For more information, see Appendix 9, which further illustrates and compares in tables the learners' performance.

5.2 shows, the percentages of acceptance of the items with null embedded subjects controlled by sentence-local antecedents for the L1 Finnish subgroups were 48.8%, 12.1%, and 2.5%; however, their percentages of acceptance were 30.4%, 9.8%, and 1.4%, respectively, for the lower-intermediate, upper-intermediate, and advanced learners for items with non-local antecedents. As for the L1 French subgroups, the percentages were 51.6%, 11.3%, and 2.0% for the former type of the sentences and 36.4%, 13.2%, and 1% for the latter.

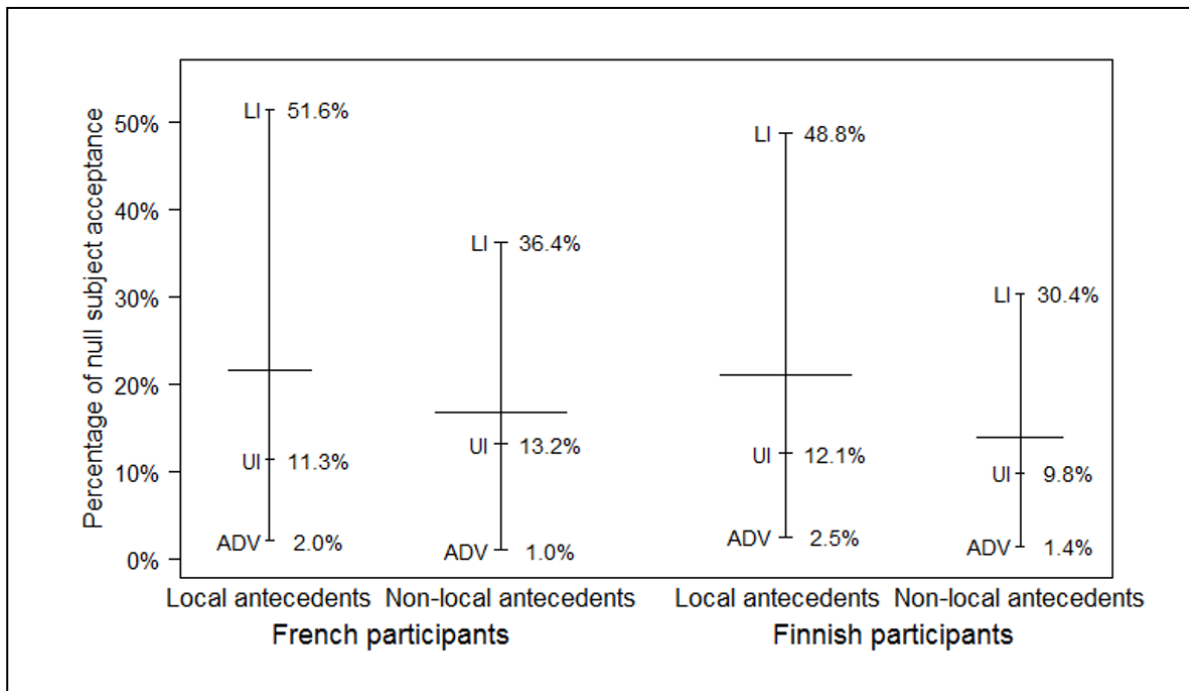


Figure 5-2. Percentages of null subject acceptances with local antecedents vs. non-local antecedents.

The differences between the performance at judging the two types of sentences within each subgroup of learners were all significant determined by the t-tests for the Finnish lower-intermediate learners ($P\text{-value} = 0.001 < 0.05$) and the French lower-intermediate learners ($P\text{-value} = 0.000 < 0.05$), and by Wilcoxon tests for the Finnish upper-intermediate learners ($P\text{-value} = 0.000 < 0.05$), the French upper-intermediate learners ($P\text{-value} = 0.000 < 0.05$), the Finnish advanced learners ($P\text{-value} = 0.0000 < 0.05$), and the French advanced learners ($P\text{-value} = 0.022 < 0.05$). However, when the Finnish- and the French-speaking learners' subgroups' results on these items' types were submitted for intergroup comparisons, the results of inferential tests conducted (as shown in Table 5.8) reveal that there were not, in fact, significant differences among subgroups of the same proficiency level.

Table 5-8. Inferential comparisons between the Finnish learners and the French learners on embedded null subjects with local antecedents and with non-local antecedents

Antecedent Position	Proficiency level	Inferential test run	P-value	L1 Finnish vs. L1 French
Null subjects with local antecedents	LI	Tukey HSD	0.916>0.05	n/s*
	UI	Tukey HSD	0.931>0.05	n/s
	ADV	Wilcoxon	0.211>0.05	n/s
Null subjects with non-local antecedents	LI	Tukey HSD	0.956>0.05	n/s
	UI	Wilcoxon	0.433>0.05	n/s
	ADV	Wilcoxon	0.716>0.05	n/s

*n/s (not significant)

It can be concluded from the above discussion that both the French and the Finnish learners differentiated in their acceptance of sentences with null pronouns depending on the position of their relevant referential antecedents; they accepted significantly more null subjects with local antecedents than with non-local antecedents. In fact, this sort of behaviour is the expected performance for both discourse-related and syntax-related reasons.

i. Discourse-related

It can be predicted, based on the results shown in section 5.4, which indicated that discourse plays a role in licensing and identification of the null elements, that the referent of an omitted argument is usually easily inferable if it is already established in the immediately preceding discourse within the same clause or the same sentence and, therefore, that the pronoun is more likely to be accepted as null compared to the null pronoun whose referent is established in broader discourse context external to the sentence. Logically, potential ambiguity, where two or more referents may be semantically suitable antecedents, tends to arise more when the search space for the appropriate referent is wider in the discourse; hence a pronoun is expected to be realized overtly to avoid such ambiguity at the discourse level.¹¹⁷ Evidence supporting the argument that discourse affects the distribution of referential null/overt

¹¹⁷ Note: this is only relevant for third-person referents; first- and second-person referents are always unambiguously identified. For more detail, see subsection 5.4 below.

subjects comes from Serratrice and Sorace's (2003) study, in which they noticed that both monolingual and bilingual Italian child-learners tend to use overt pronominal forms when the referent was not easily accessible (i.e., far back in the discourse) or new to discourse.

It should be mentioned here that the results presented above in Figure 5.2 and in Table 5:8 provide evidence that discourse and discourse-pragmatic factors play an important role in licensing the presence of null elements in L2A. (This will be discussed in detail in section 5.4 below.) However, it is expected that such discourse roles operate within the boundaries imposed by the grammar (see Serratrice and Sorace, 2003; Serratrice, 2005; Frascarelli, 2007).¹¹⁸ This leads us to the syntax-related reason.

ii. **Syntax-related:**

It can be anticipated, if learners indicated that null subjects can take non-local antecedents crossing the clause or the sentence boundary, that null subjects whose referential antecedents are clause-bound would have higher chances of being accepted as null, in accordance with Chomsky's (1995) Minimal Link Condition (1995) and Richards' (2001) Attract Closest Condition. In the context of pro-drop grammar, these conditions assume that embedded null subjects would attract or prefer the closest suitable antecedents over distant suitable ones.¹¹⁹

It can be concluded based on the French performance (illustrated by Figure 5.2 above) the French participants' performance illustrated in Table 5.1 above was not due to an uncontrolled variable. In other words, as their performance obeys the syntactic conditions of the Minimal Link Condition or the Attract Closest Condition as

¹¹⁸ It is well documented that the interaction between syntax and discourse pragmatics affects the distributions of referential null/overt subjects in pro-drop languages and in child production of the null/overt subjects (see, among many others, Frascarelli, 2007; Heycock and Filiaci, 2004; Montrul, 2004; Orsolini and Di Giacinto, 1996; Orsolini, Rossi, and Pontecorvo, 1996; and Tsimpli, Sorace, and Sorace, 2004).

¹¹⁹ See also the Economy Condition (Chomsky, 1989) in which these two conditions are related. For more information, see Radford (2004).

well as the discourse factors, results which show that the French learners of English accept significantly more null subjects with local antecedents compared to null subjects with non-local antecedents tallies with the results that the French-speaking participants accepted ungrammatical sentences with null phonetic spellout subjects—a conclusion based on which, in addition to the findings presented in Figure 5.2, hypothesis H1 is rejected.

For comparative and contrastive purposes, the Finnish results, presented in Figure 5.2 above, will be discussed later in this section after presented the results of the L1 Arabic participants with respect to their acceptance of referential embedded null subjects with local antecedents vs. non-local antecedents.

As for the Arabic-speaking participants, their results showed a reverse performance pattern in comparison to their Finnish and French peers as subgroups; compare Figure 5.3 below with Figure 5.2 above.

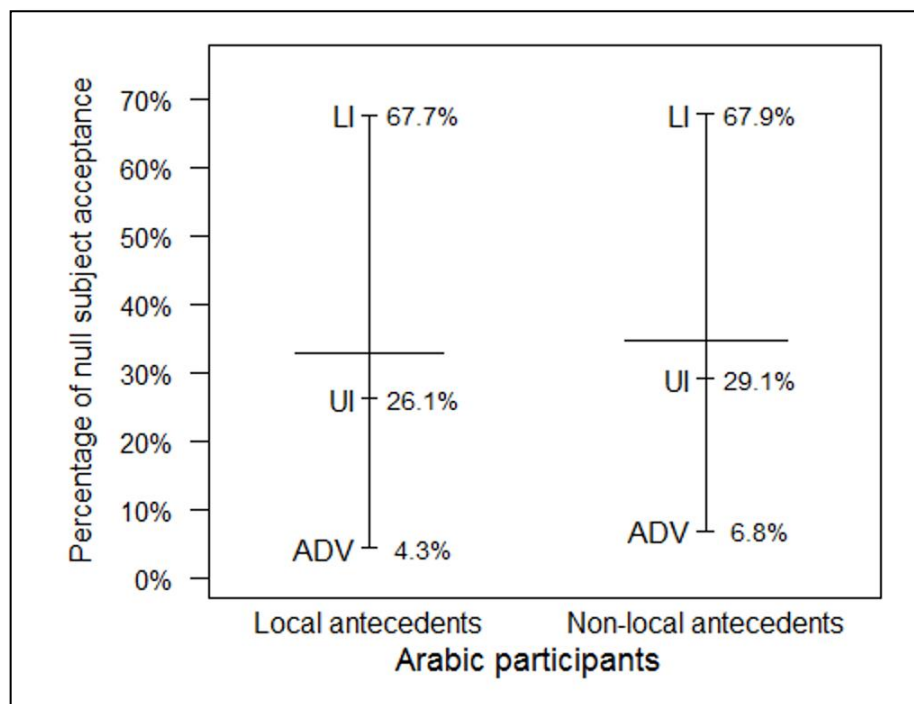


Figure 5-3. Percentages of null subject acceptances with local antecedents vs. non-local antecedents.

As Figure 5.3 shows, the percentages of acceptance of the items with null embedded subjects controlled by sentence-local antecedents for the lower-intermediate, upper-intermediate, and advanced learners were 67.7%, 26.1%, and

4.3%, respectively, whereas their percentages of acceptance were 67.9%, 29.1%, and 6.8%, respectively, for the lower-intermediate, upper-intermediate, and advanced learners for items with non-local antecedents. Although it seems from the percentages that the Arabic participants at all proficiency levels accepted marginally more embedded null subject sentences with non-local antecedents than those with local antecedents, the inferential tests run specifically on the lower-intermediate and upper-intermediate subgroups indicate that these learners' judgments within each subgroup for the null subject sentences with sentential external referents were significantly different from the ones with sentential internal referents; the result obtained by the t-test for the lower-intermediate Arabic subgroup was $P\text{-value} = 0.000 < 0.05$; the result arrived at via the Wilcoxon test for the upper-intermediate Arabic subgroup was $P\text{-value} = 0.001 < 0.05$. However, the result arrived at via the Wilcoxon test for the advanced participants shows that there were no significant differences between their performance on these two types of items ($P\text{-value} = 0.134 > 0.05$). Such differences in the performance between the lower-intermediate and upper-intermediate subgroups on the one hand and the advanced subgroup on the other hand can be attributed to the proficiency increase in English with time.

Given how pro-drop works in Arabic where external antecedents are allowed, as opposed to Finnish, where pro-drop antecedents have to be local and preferably no further away than the previous clause, the difference between the Arabic participants and the Finnish participants, who accept significantly more null subjects with local antecedents compared to null subjects with non-local antecedents, is expected and in agreement with the Modulated Structure Building approach (Hawkins, 2001) which allows for influence of the L1 during the subsequent process of structure building.¹²⁰ Such a difference in performance supports the distinction

¹²⁰ As a native speaker of Arabic, I feel that, when there are two potential antecedents that an embedded third-person null subject can corefer with due to the absence of enough discourse or extra-linguistic context, the null subject prefers the non-local referent; so, it could be argued that this suggests that the clause-external option for licensing a null subject transfers. However, it should be mentioned that the situation is more complex than what is stated here; the choice of the referential antecedent truly depends on several factors, including the realisation of its potential referents (overt vs.

between types of pro-drop discussed in Chapter 2 and also provides evidence for Holmberg's (2010a) argument that there are several null subject parameters. Holmberg (2010a) asserted, "It is widely recognized that null subjects can be derived in more than one way, and that, therefore, more than one parameter is involved determining whether subject pronouns can be null or not in a given language" (p. 88). This theoretical argument, when linked to the findings presented above in Figures 5.2 and 5.3, raises the question: Are there several null subject parameters in SLA? Finding answers to this question will help researchers further explore and understand the issue of null subject transfer in SLA.

5.2.2 The translation task

5.2.2.1 The English IL of the French, Finnish and Arabic speakers: Basic descriptive and inferential statistics

i. The English ILs: French speakers

The descriptive statistics for the data obtained by the translation task for all three subgroups of French-speaking English L2 learners are presented below in Table 5.9. It reveals that the advanced learners as a subgroup provided no subjectless sentences when translating the French sentences into English; the mean of sentences with embedded null subjects produced by this subgroup of learners was 0.00 (StdDev 0.00). The upper-intermediate and lower-intermediate learners, as subgroups, were not as accurate in producing the grammatically correct English equivalent sentences compared to the advanced subgroup. As Table 5.9 shows, both subgroups performed comparably; the mean score of omission for the upper-intermediate subgroup was 0.10 (StdDev 0.40) and 0.11 (StdDev 0.33) for the lower-intermediate subgroup.

null) and its potential referent grammatical category (pronoun vs. noun). Future research is necessary to investigate how the thematic roles of the possible referents influence referent accessibility and choice of the ambiguous pronoun in Arabic. I am not aware of any study that has been done at this point to investigate this issue in Arabic.

Table 5-9. Descriptive statistics for ungrammatical sentences with missing subjects produced by French-speaking learners

Proficiency level	Included translation	No. of omission	Max	Min	Mean	StdDev	Percentage of omission
ADV (N= 27)	405	0	0	0	0.00	0.00	0.0
UI (N= 31)	464	3	2	0	0.10	0.40	0.7
LI (N= 09)	135	1	1	0	0.11	0.33	0.7

To see if the learners' proficiency levels have affected their performance with reference to the use of null referential embedded subjects in English, I performed Kruskal Wallis inferential tests on the results. The test results showed that there is no statistical significant difference among these proficiency subgroups (P-value = $0.300 > 0.05$). The test results, in other words, indicated no relationship between proficiency and performance. To double check this finding, a series of non-paired Wilcoxon tests were run to compare the results of two samples at a time; the differences between the lower-intermediate and the upper-intermediate subgroups, between the lower-intermediate and the advanced subgroups, and between the upper-intermediate and the advanced subgroups were all not significant, with P-value = $0.697 > 0.05$, P-value = $0.095 > 0.05$, and P-value = $.191 > 0.05$, respectively. This confirms that the learners' translation accuracy rates with respect to the use of embedded overt subject pronouns do not change with increased proficiency in English.

To see how the L2 learners diverge from the native speakers and at the same time whether they converge to become native-like, we need to compare their results against one another (native language vs. interlanguage). However, because it goes without saying that the English native speakers control group did not complete the translation task, it is assumed that they dropped none of the subject pronouns based on the fact that they accepted none of the ungrammatical sentences with null subjects (see Final Piloting below in subsection 4.3.3.2).

When inter-subgroup comparisons were conducted for the purpose of comparing the results of each proficiency subgroup presented in Table 5.9 above with the hypothesised results of the English native controls illustrated in appendix 10, the differences found between the lower-intermediate and the native speakers and between upper-intermediate subgroup and the native speakers were not significant (P-value = $0.450 > 0.05$ and P-value = $0.527 > 0.05$, respectively, for the lower-

intermediate subgroup and upper-intermediate subgroup); the advanced learners behaved completely like native English speakers by producing no null subject clauses at all. This suggests that all French participants, regardless of their proficiency levels, behaved like native English speakers with respect to producing no null subject clauses.

The translation task results of the French learners of English, presented in this section, suggest that the L1 setting transfers in L2A; the French participants, regardless of their proficiency levels, behaved like native English speakers with respect to producing no null subject clauses. This result will be discussed later in detail in subsections 5.2.2.2 and 5.2.2.3, where all of the subgroups' results (L1 French-speakers, L1 Finnish-speakers, and L1 Arabic-speakers) from the translation task are compared and discussed.

ii. The English ILs: Finnish speakers

The Finnish-speaking learners showed similar acquisition pattern to that of their French-speaking counterparts as far as embedded subjects are concerned. As the results in Table 5.11 indicate below, the lower-intermediate Finnish subgroup overall dropped more subjects with a mean of 0.83 (StdDev 2.04) when translating sentences into English than did their upper-intermediate peers as a subgroup, with a mean of 0.18 (StdDev 0.53). The advanced learners managed to translate all the given Finnish sentences with null embedded subjects into the correct grammatically equivalent English sentences with overt embedded pronominal forms behaving exactly like English native speakers (mean of subject omission = 0.00).

Table 5-10. Descriptive statistics for the ungrammatical sentences with null subjects produced by Finnish-speaking learners

Proficiency level	Included translation	No. of omission	Max	Min	Mean	Std Dev	Percentage of omission
LI (N= 06)	79	5	5	0	0.83	2.04	6.3
UI (N= 17)	254	3	2	0	0.18	0.53	1.2
ADV (N= 53)	787	0	0	0	0.00	0.00	0.0

To see the impact of the proficiency levels on the test-takers' performances, a subgroups comparison of pronouns omission was conducted using a Kruskal Wallis

test. The test revealed that there was marginal significant difference in use of pronouns among the three subgroups of the Finnish L2 English learners (P-value = $0.024 < 0.05$). The fact that the advanced learners behaved completely like native English speakers by producing no null subject clauses at all may indicate that this subgroup significantly outperformed the lower-intermediate and upper-intermediate subgroups. This observation is supported statistically; the differences between the advanced and the lower-intermediate subgroups and between the advanced and the upper-intermediate subgroups registered by Wilcoxon tests showed significance levels of P-value = $0.000 < 0.05$ and P-value = $0.013 < 0.05$, respectively. However, the difference between the lower-intermediate and upper-intermediate subgroups did not reach statistical significance by the same inferential test (P-value = $0.244 > 0.05$). This implies that these learners' abilities to translate sentences into L2 English appropriately increase as their proficiency in English increases.

However, despite this marginal significant difference in use of pronouns existing among the three proficiency subgroups, when the Finnish participants' subgroup's results were statistically compared with the results' of the controls (see appendix 10), the differences between the lower-intermediate subgroup and the native control group and between the upper-intermediate subgroup and the control group were both not significant by Wilcoxon tests (P-value = $0.140 > 0.05$ and P-value = $0.389 > 0.05$, respectively). These results indicate that both subgroups performed within the range of the native control performance - These results indicate that both subgroups performed within the range of the native control performance,- which may suggest that the L1 setting (partial pro-drop) does not transfer at the early stages of L2 development, but note the L2 initial-state IL is not currently under investigation (see the discussion in subsections 5.2.1.2 and 5.2.1.3).

iii. The English ILs: Arabic speakers

The Arabic-speaking learners showed a different developmental pattern from that of their French- and Finnish-speaking counterparts. As the results in Table 5.12 below indicate, the lower-intermediate Arabic subgroup overall produced more sentences with null subjects with a mean score of 1.75 (StdDev 2.62) than did the upper-intermediate subgroup with a mean score of 0.64 (StdDev 1.65), who themselves (the upper-intermediate) dropped more subject than did their advanced peers as

a subgroup with the mean of 0.18 (StdDev 0.40). These means were submitted to a Kruskal Wallis test to see if there was a relationship between performance and English proficiency. The results indicated that there was a significant difference among the three proficiency subgroups with respect to embedded subject drop (P-value = 0.014<0.05). In particular, statistically significant differences existed only between the lower-intermediate and the upper-intermediate subgroups and between the lower-intermediate and the advanced subgroups under the Wilcoxon tests, with P-value = 0.025<0.05 and P-value = 0.017<0.05, respectively. The difference between the upper-intermediate and the advanced subgroups did not achieved significance under this test (P-value = 0.966>0.05). This indicates that performance becomes better as proficiency in English increases with regard to the variable under the analysis.

Table 5-11. Descriptive statistics for ungrammatical sentences with null subjects produced by Arabic participants

Proficiency level	Included translation	No. of omission	Max	Min	Mean	StdDev	Percentage of omission
LI (N= 16)	224	28	9	0	1.75	2.62	12.5
UI (N= 14)	173	9	5	0	0.64	1.65	5.2
ADV (N= 11)	165	2	1	0	0.18	0.40	1.2

The Arabic participants differed significantly from the native speakers at only the lower-intermediate level as far as embedded subjects are concerned, according to the Wilcoxon test (P-value = 0.010<0.05), which suggests that the L1 setting (pro-drop) transfers at the initial stages of L2A (see the discussion in subsections 5.2.1.2 and 5.2.1.3). However, the differences between the upper-intermediate subgroup and the native group and between the advanced subgroup and the native control group did not achieve significance under this test (the Wilcoxon), at P-value = 0.341>0.05, and P-value = 0.280>0.05, respectively.

5.2.2.2 Comparisons across all language subgroups: ILs vs. ILs of the same proficiency level

The results of the inferential tests conducted in the previous subsection showed that there was no relationship between proficiency levels and performance among the French-speaking learners in contrast to their Finnish-speaking and Arabic-speaking

peers, whose inferential tests results indicated that there were significant differences among proficiency subgroups of the same L1 language with respect to embedded subject drop. Yet, based on the results of P-values arrived at via the Kruskal Wallis tests, the significant differences between proficiency and performance were clearly greater among the lower-intermediate, upper-intermediate, and advanced Arabic speakers ($P\text{-value} = 0.014 < 0.05$) than among their Finnish counterparts ($P\text{-value} = 0.024 < 0.05$). These results may suggest that learners from different L1 backgrounds approach the task of translation differently as far as embedded subjects are concerned. However, because all groups of learners regardless of their L1 linguistic backgrounds were generally successful in producing grammatically correct English sentences, as indicated by the low mean scores of subject omissions illustrated in Tables 5.9, 5.11, and 5.12 above (represented graphically in Figure 5.4 below), intergroup comparisons among the L1 French-speakers, L1 Finnish-speakers, and L1 Arabic-speakers of the same proficiency level are required before these findings can be attributed to transfer from the L1.

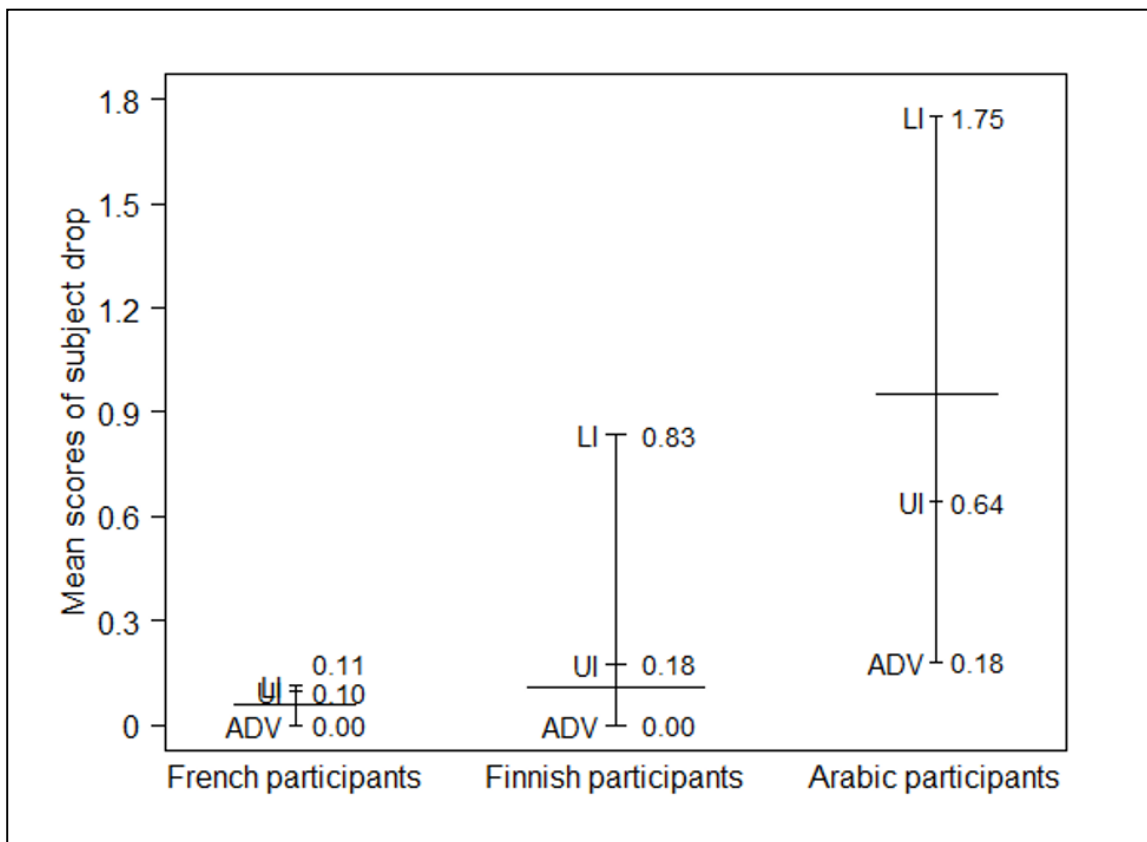


Figure 5-4. Intergroup comparisons between the subgroups of learners of the same proficiency level.

Therefore, a series of Kruskal Wallis tests were run for these comparative and contrastive purposes. The statistical data revealed significant differences among the three French, Finnish, and Arabic lower-intermediate-level subgroups in terms of the learners' performances with respect to their use of English-embedded subject pronouns ($P\text{-value} = 0.027 < 0.05$), as well as between three advanced subgroups of learners ($P\text{-value} = 0.001 < 0.05$). However, this test showed no significant difference among the three upper-intermediate-level subgroups of learners ($P\text{-value} = 0.623 > 0.05$). This inferential result of the three upper-intermediate-level subgroups was supported by the results of the Wilcoxon tests, which compared the performance of two samples at the same time; the differences between the French upper-intermediate and the Finnish upper-intermediate subgroups, between the French upper-intermediate and the Arabic upper-intermediate subgroups, and between the Finnish upper-intermediate and the Arabic upper-intermediate subgroups were all not significant, at $P\text{-value} = 0.544 > 0.05$, $P\text{-value} = 0.358 > 0.05$, and $P\text{-value} = 0.759 > 0.05$, respectively. As for the lower-intermediate-level subgroups, only the French subgroup was significantly different from the Arabic subgroup, as the results obtained via the Wilcoxon test indicated ($P\text{-value} = 0.014 < 0.05$); the differences between the French and Finnish subgroups and between the Finnish and Arabic subgroups did not achieve significance under this test ($P\text{-value} = 0.311 > 0.05$ and $P\text{-value} = 0.309 > 0.05$, respectively). However, the performance of the Arabic advanced subgroup was significantly distinct from both that of the French advanced subgroup and the Finnish advanced subgroup by Wilcoxon tests ($P\text{-value} = 0.027 < 0.05$ and $P\text{-value} = 0.002 < 0.05$, respectively). The French advanced and Finnish advanced subgroups performed similarly.

Despite the fact that both the Finnish and Arabic lower-intermediate subgroups and the Finnish and French lower-intermediate subgroups performed similarly, it could be argued that, because the performance of the French lower-intermediate subgroup was distinct from the Arabic lower-intermediate subgroup, the differences and even the similarities among the performances of the French-, Finnish-, and Arabic-speaking L2 learners of English on this translation task were the product of L1 transfer. As transfer is mainly associated with the early stages of L2 grammatical development, and its effects could last for a while, this argument is supported by the statistical differences that exist between the performances of the Arabic and Finnish

advanced subgroups, which indicates that Finnish speakers converged to native-like usage faster than the Arabic speakers, even though both groups dropped subjects similarly at the lower-intermediate level. If the L1 had no effect, one would expect all three groups to behave similarly at both the initial and final stages of development. However, one should remain sceptical about drawing firm conclusions about whether null subject parameter settings transfer in L2A based on the translation task results.

5.2.2.3 Discussion of the translation task's results

The results suggest that the learners from different linguistic backgrounds vary with regard to the production of referential embedded subject pronouns in English. The French participants behaved completely like native English speakers from early on as a direct consequence of having an L1 that does not allow subject omission. Like their French peers, the Finnish participants performed within the native-like ranges at the lower-intermediate-level, but there was nevertheless a marginal statistical significant difference in their abilities as proficiency subgroups to translate sentences appropriately, unlike their French counterparts, whose abilities to perform the task perfectly were constant from the lower-intermediate level onwards. This result indicates that Finnish speakers showed a tendency to omit referential subjects more than their French counterparts, although not at a statistically significant level. Alternatively, such a tendency was statistically confirmed when the results of the Finnish lower-intermediate subgroup were compared with the Arabic lower-intermediate subgroup; there was no statistical difference between their performance. Such successful performance can be attributed to the fact that Finnish optionally allows null subjects.¹²¹ The Arabic participants, however, produced null subject embedded clauses at only the lower-intermediate level as a result of having an L1 where such omissions are required in most cases (see Chapter 2), but nevertheless they managed to converge to native-like performance at the upper-intermediate level. Therefore, these results seem to confirm the predictions of hypothesis H1 that were formulated based on the principles of the FT/FA hypothesis,

¹²¹ The fact that colloquial Finnish is a non-null subject language, where subject omission does not occur in any persons, may have helped these subjects to perform native-like (for more detail about colloquial Finnish, see Vainikka and Levy, 1999).

especially if we take into consideration the fact that the L2 initial-state IL is not currently under investigation in the present study; the French participants were found not to drop subjects in English because English adopts the same [-pro-drop] parametric value that French does. On the other hand, the Finnish participants were found to drop a lower number of subjects in English than their Arabic counterparts which can be seen as a direct consequence of having an L1 that allows null subjects but under more restricted conditions than in Arabic.¹²²

The results of the present study seem to support the findings of previous studies that reported that subject drops are attested in the productions of adult native speakers of null subject languages learning non-null subject languages, especially in the initial stages of L2 development (Hilles, 1986; LaFond, 2001; Licerias, 1988; Licerias and Díaz, 1999; Phinney, 1987; Sauter, 2002). Because the L2 initial-state IL is not currently under investigation, the subject drops observed in the production of the lower-intermediate-level Arabic and Finnish participants, despite the low rates, can be explained by the argument that dropping rates of pronominal subjects in L2A are not similar in matrix clauses or in embedded clauses. For example, Phinney (1987, p. 234) observed that “most of the omissions occurred in subordinate or conjoined clauses where the discourse was already focused on the subject.”¹²³ Note that only subject omissions occurring in subordinate clauses are at the centre of attention in the present study.

However, the present results differ from Orfitelli and Grüter (2013), who found that L1 Spanish learners of English showed no evidence of subject drop, even in early stages of L2 development. However, this apparent contradiction can be explained by the argument that Orfitelli and Grüter’s study was not successful at creating appropriate contexts for subject pronominalization; for more detail, refer to Chapter 3 where this study was critically reviewed.

¹²² The fact that colloquial Finnish is a non-null subject language may also have helped these the Finnish participants.

¹²³ See Licerias and Díaz (1999) for a more detailed argument about dropping constraints.

The results of the translation task are not going to be analysed based on the position of the potential referential antecedents of the embedded missing subjects due to the fact that subject drop was generally scarce in the participants' production, unlike their acceptances of null subjects in the GJ task, which were frequent. Addressing this asymmetry reflected by the results between subject drop on the production task and null subject acceptance on the GJ task will be the goal of the following section.

5.2.3 The grammatical intuition data vs. translational production data: Comparisons within each subgroup

Thus far, the results that have emerged from the production and receptive tasks, as presented in the previous subsections have provided conflicting conclusions about whether null subject parameter settings transfer in L2A. The results from the GJ task show evidence that all learners, learners regardless of their linguistic backgrounds, start off with pro-drop as their initial grammar lacks functional projections, whereas the results from the translation task suggest that the L1 setting transfers in L2A. Furthermore, a comparison of Figures 5.1 and 5.4 suggests that the learners' performance varies from task to task; that is, the results show a clear trend indicating the learners' performance in judging ungrammatical structures was poorer than their performance in the translation task with respect to the proper use of referential subject pronouns, regardless of their L1 linguistic backgrounds or their proficiency level.

The goal of this subsection is to look in depth at the learner subgroups' results to compare the overall performance in the translation task to that in the GJ task within each subgroup. To achieve this goal, the quantitative differences or the degree of variability regarding the L2 non-target-like performance on null subject pronouns across the two tasks are examined within each subgroup in terms of percentages of pronoun misuse (subject omission vs. null subject acceptance).

5.2.3.1 Comparisons on subjects' null realisation across the two tasks

i. IL variability within each lower-intermediate-level subgroup

As Figure 5.5 below shows, the three lower-intermediate subgroups provided sentences with null embedded subjects only at percentage rates of 0.7%, 6.3%, and 12.5% for the French, Finnish and Arabic subgroups of learners, respectively; in judging ungrammatical sentences with missing embedded pronominal subjects, the same subgroups respectively accepted 49.3%, 46.1%, and 67.8% of the mistakes.

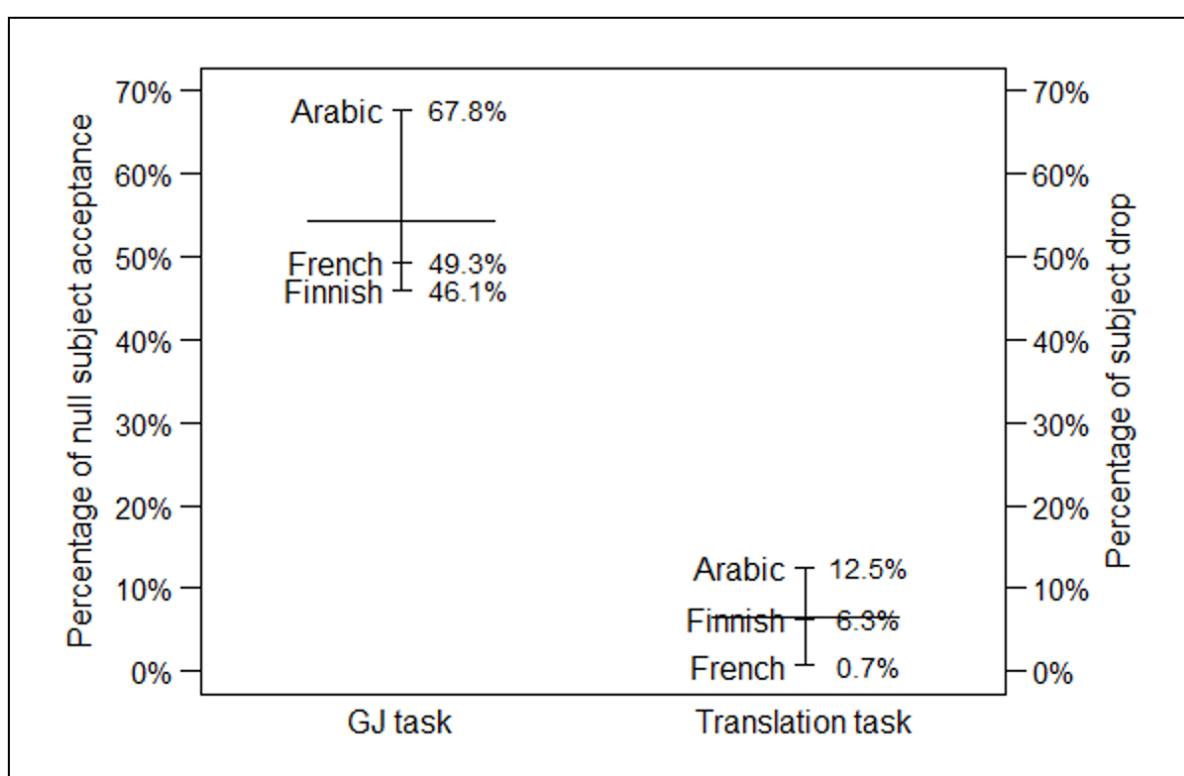


Figure 5-5. Pairwise comparisons between the translation task and the GJ task performances within each lower-intermediate-level subgroup of learners.

The differences among these descriptive results suggest that learners approached the two tasks differently, in the sense that their performances vary from task to task. This suggestion was proved for all of the lower-intermediate subgroups through the results of inferential tests that confirmed there were significant differences in the lower-intermediate learners' performances within each subgroup on non-target-like null subject pronouns across the GJ task and the translation task. The results arrived at via the Wilcoxon tests were $P\text{-value} = 0.009 < 0.05$ for the French lower-intermediate learners, $P\text{-value} = 0.036 < 0.05$ for the Finnish lower-

intermediate learners, and $P\text{-value} = 0.000 < 0.05$ for the Arabic lower-intermediate subgroup.

ii. IL variability within each upper-intermediate-level subgroup

The same trend that subject drops were lower in the learners' production can be observed by examining the results obtained by the upper-intermediate-level subgroups found in Figure 5.6 below; the learners produced far fewer sentences with null subjects as opposed to the ungrammatical sentences with null subject they accepted (0.6%, 1.2%, and 5.2% vs. 11.6%, 11.9%, and 26.6% for the French, Finnish, and Arabic subgroups, respectively). Accordingly, the differences in performance within each upper-intermediate subgroup on non-target-like null subject pronouns across the GJ task and the translation task were all shown to be significant by Wilcoxon tests ($P\text{-value} = 0.000 < 0.05$, $P\text{-value} = 0.001 < 0.05$, and $P\text{-value} = 0.001 < 0.05$ for the French, Finnish, and Arabic upper-intermediate subgroups, respectively).

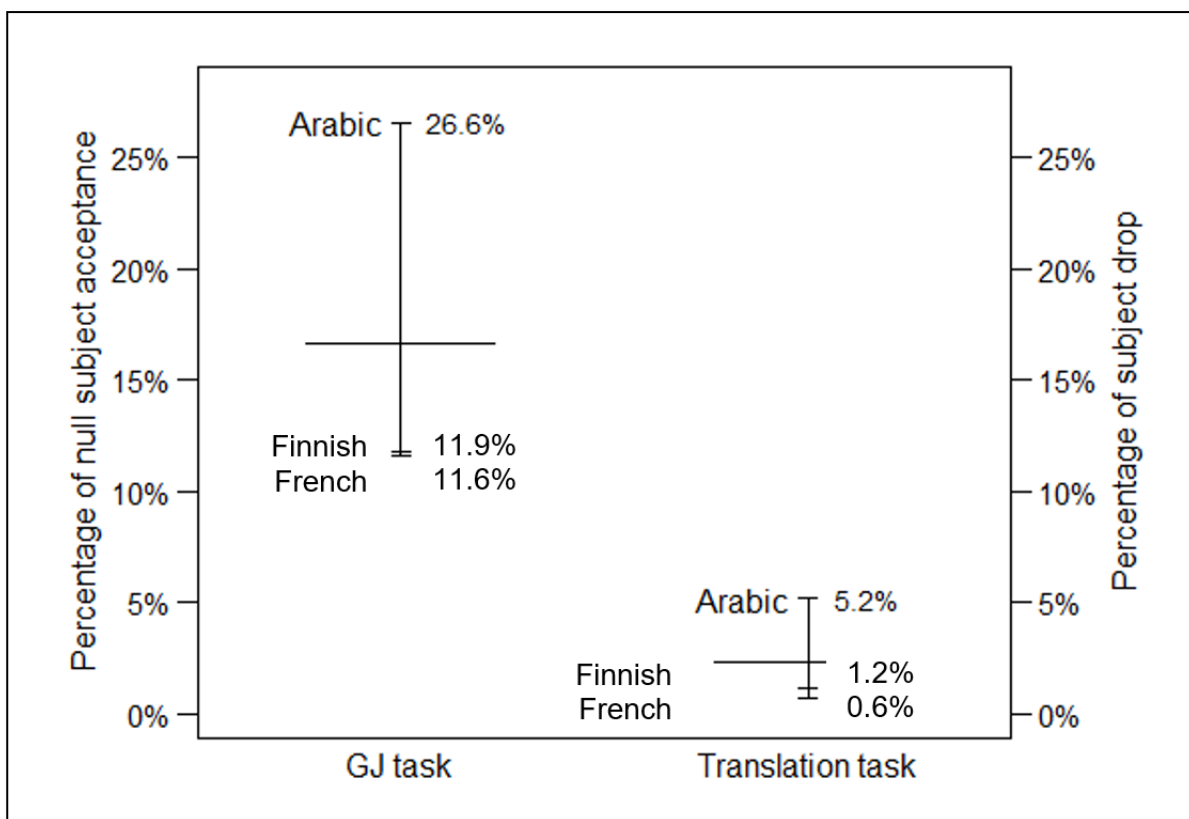


Figure 5-6. Pairwise comparisons between the translation task and the GJ task performances within each upper-intermediate-level subgroup of learners.

iii. IL variability within each advanced subgroup

This trend holds throughout all the advanced subgroups as well. All three subgroups showed high success rates at rejecting items with null embedded subjects, but they performed with very high accuracy rates at producing grammatically correct English sentences with overt embedded subjects. As Figure 5.7 shows, subject pronoun omission occurred in 0.0% of the cases for both the French and Finnish advanced subgroups, while subjectless clauses were accepted 1.9% and 2.3% of the times for the same subgroups. In the Arabic-speaking advanced subgroup, subject omission occurred in 2.4% of the cases, whereas sentences with null embedded subjects were accepted 4.7% of the time. The differences in performance within each subgroup between dropping embedded subjects and accepting embedded null subjects are all highly significant statistically as determined by Wilcoxon tests for the three advanced-level French-, Finnish-, and Arabic-speaking subgroups (P-value = 0.021<0.05, P-value = 0.000<0.05, and P-value = 0.034<0.05, respectively).

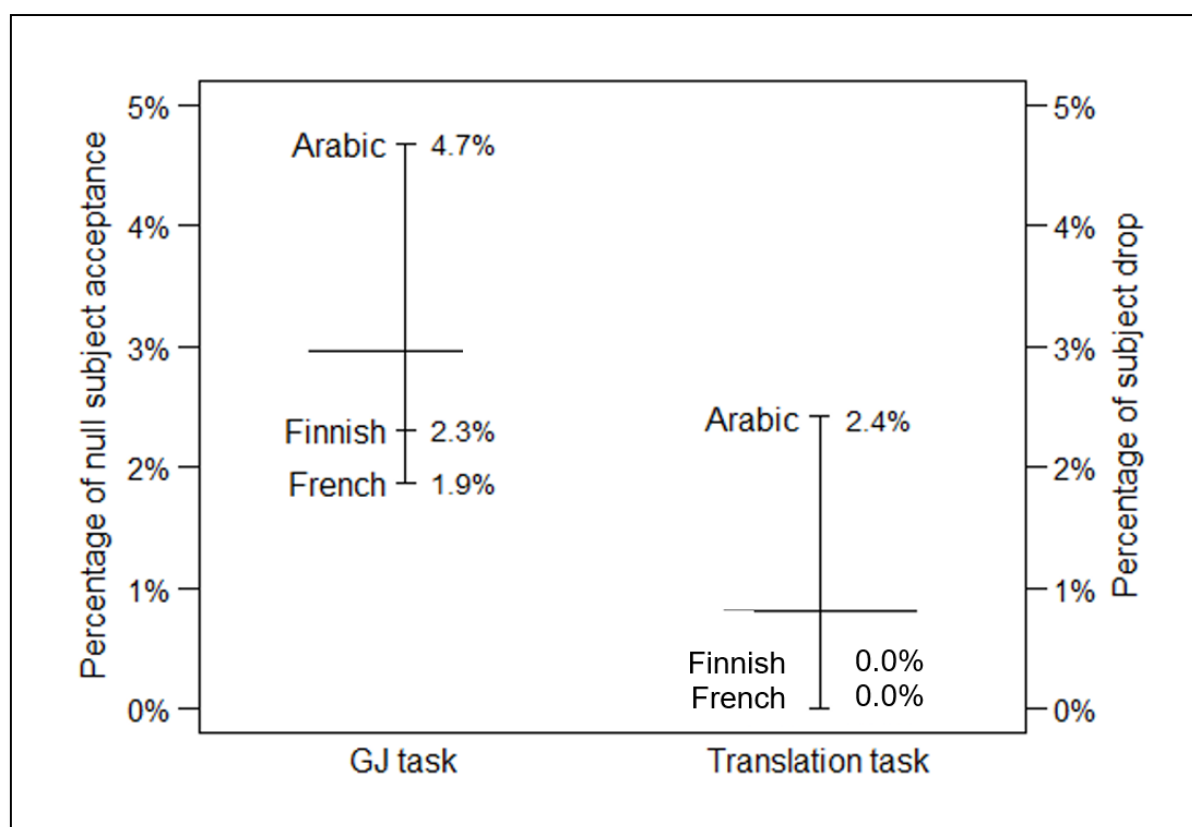


Figure 5-7. Pairwise comparisons between the translation task and the GJ task performances within each advanced subgroup of learners.

5.2.3.2 Discussion of results divergence/variability across the two tasks

The discussion in the previous subsection made it evident that participants in proficiency subgroups, regardless of their L1 backgrounds, did not perform consistently across different task types. Even though their performance generally increased as their proficiency increased, the GJ task seemed to be much more difficult for L2 learners of English compared to the translation task, which appeared to be quite easy as far as embedded subjects are concerned. This contrast emerged when the results of the GJ task were compared to the results of the production task and disconfirm hypothesis H2 which predicted that all the learners, within each subgroup, would perform consistently across the two task types.

Generally speaking, the contrast within the performances of the learners appears consistent with the findings of previous empirical studies when considering the differences in the rates reported between subject omissions and null subject acceptances.¹²⁴ These studies indicate that [+prodrop] learners acquiring a [-prodrop] language consistently accept ungrammatical sentences with referential null subjects with approximate rates of 24–41% (Davies, 1996; Judy, 2011; Judy and Rothman, 2010; Orfitelli and Grüter, 2013; White, 1985, 1986), whereas such learners are reported to drop subjects when tested with production tasks only at the earliest stages of L2 development, if it exists, at approximate rates of 0–13% (Hilles, 1996; Phinney, 1987; Orfitelli and Grüter, 2013).

The question that logically follows is this: What are the possible causes of such inconsistent use of subject pronouns (overt vs. null) noticed in the participants' performance across both tasks (the intuitive and the translation tasks)? In other words, why do the learners persistently accept referential null subjects in the GJ task beyond the stage of L2 development when they have established the requirement for overt subjects in their production?

¹²⁴ Notably, the results of several studies conducted to investigate the L2 learning of a wide variety of linguistic structures lend support to the observation that the use of different elicitation tasks would yield differences in the performance of L2 learners (Lee, 2014; Tarone, 1985; Wright, 2010)

A possible explanation to account for this sort of inconsistency is to argue that [+prodrop] is part of the learners' IL grammars of English, so that their ILs permit both null and overt referential subjects (this will be explained below more clearly when considering the notion of judgment strength of preferences). However, the written nature of the translation task prompted the learners to strongly prefer using overt pronouns over null forms, but when they are forced to make judgements about given sentences, null subjects may surface in their acceptability due to the fact that acceptability is a gradient concept.¹²⁵

To investigate this argument in some depth, the participants' strength of preferences for the overt pronouns over null forms need to be measured. One way to do so here (and recall what was discussed in relation to the rating scale used in the GJ task, how it worked, and how the data obtained were analysed in Chapter 4) is by looking at the results of the doubtful category (*possibly incorrect option*) only when the right corrections were provided; namely, only the PIT scores (refer to subsection 4.3.6.1: The GJ Task: The Adopted Marking Method). Figure 5:8 below displays the percentages of the PIT scores of the total rejected ungrammatical sentences with null subjects, which are made of the PIT scores added to the CIT scores (clearly incorrect—the true correction was provided):¹²⁶

¹²⁵ See section 4.3.2.1.2 for more detail about the differences between grammaticality and acceptability.

¹²⁶ For more details, see to subsection 4.3.6.1: The GJ Task: The Adopted Marking Method.

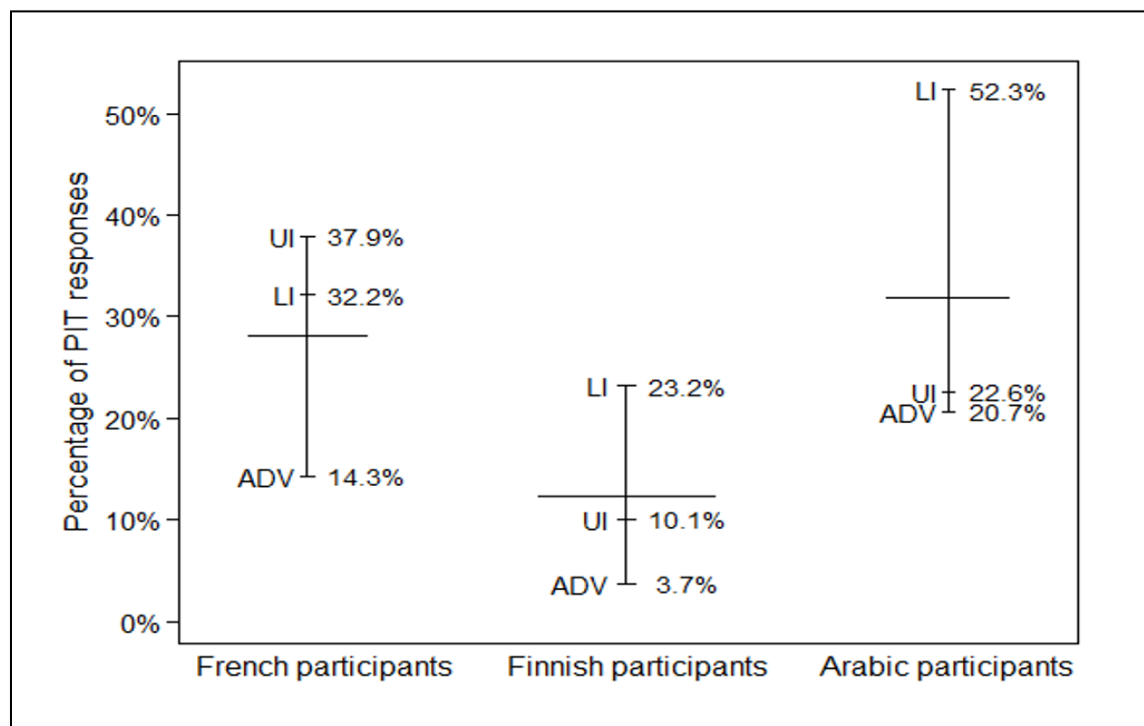


Figure 5-8. The subgroups' overall percentages of the PIT scores to ungrammatical sentences with null subjects.

The above figure illustrates the participants' strength of preferences for overt pronouns over null forms; it shows that the French lower-intermediate subgroup participants, for example, by choosing the *possibly incorrect* option were reluctant/hesitant concerning 32.17% of the ungrammatical items they rejected. They have an inkling that these items could be correct, even though they corrected the errors that rendered the sentences ungrammatical. Similar observations can be noticed in all other subgroups of participants, regardless of the amount/size of the percentage of PIT rejection and/or the degree of doubt (little or great) when considering Figure 5.8 again.¹²⁷ Hence, these PIT percentages can be taken as possible indices of the assumptions:

- that choosing the possibly incorrect option indicates that the null subject option [+ null subject] is part of the learners' IL grammar;
- that managing to provide the right correction indicates a preference for the

¹²⁷ Due to space limitations, only the results of the PIT scores of the French lower-intermediate subgroup is presented here as an example to illustrate and support the argument.

overt pronouns over the null ones; and

- that choosing the possibly incorrect option and then managing to correct the sentence accurately support the idea that acceptability is a gradient concept.

Different explanations have been proposed to account for such intraindividual variability observed in the learner's performance on the same day across any set of two or more tasks. For example, generative linguists have explained such variability by the fact that performance is not a perfect reflection of competence on all occasions. As they are interested in "what a speaker knows about language as an internal property of human mind rather than something external [the produced utterances]" (Chomsky, 1988, p. 36), they often argue for the use of GJ tasks to ascertain learners' intuitions rather than using other tests that might not reflect the learner's IL (see the relevant dissection in Chapter 2 and 4). Psycholinguists (e.g., Orfitelli and Grüter, 2013; Wright, 2010) seem to be more interested in learner's variability even though they share the view with generative linguists that IL is a static/constant system at any given point in time; they link such variability to factors internal to the learner such as attention, processing limitation, and demands of short-term memory. In contrast, sociolinguists generally view IL as a dynamic linguistic system varying at any given point according to learner's interaction with the environment (i.e., formal vs. informal context, speaking vs. writing task), which results in noticed variability (for more detail cf. Lowie and Verspoor, 2015; Throne, 1988; Verspoor, Lowie, and van Dijk, 2008). Before leaving this discussion, it is important to comment briefly on these alternative explanations:

- i. Various criteria have been implemented in the present study to make learner's performance on both tasks reflect his/her IL to a great extent (see Chapter 4).
- ii. If variability is only caused by extragrammatical factors, null subjects observed in L2 learners' acceptances or productions must also have extragrammatical causes, not due to underlying grammars that permits them [+prodrop]. However, the findings of the present study and previous studies speak against this view; they show evidence of transfer of null subjects from

L1 to L2. I will come back to this point in more detail when the issue of null subject licensing is discussed in section 5.4.

- iii. As for the sociolinguistic view of variability, I think it is impossible for tasks to alter the status of the IL; otherwise, a learner would be identified as being at two or more different developmental levels at the same time.

At this point, I turn to the notion of parameter resetting. To recap, the empirical data presented and discussed here provide evidence that while Finnish and Arabic L2 learners of English continued to accept null subjects even at the advanced stage of L2 development (despite the fact null subjects had disappeared from their speech during the early stages of acquisition), L1 French speakers fully converged on the English native-like usage of overt subjects. These findings are open to two quite different scenarios with respect to the null subject parameter resetting in L2A: the parameter either can be reset or cannot be reset.

According to the first scenario, one can argue based on the results that because the null subject would not be fully eliminated from the learners' IL grammars—though it will continue to coexist with the overt pronominal form even though it does not surface in their productions during the intermediate and the advanced stages of L2 acquisition—that new parameter settings cannot be acquired in SLA. This sort of argument is consistent with the No Parameter Resetting Hypothesis (Hawkins and Chan, 1997; Smith and Tsimpli, 1995), which assumes L2 learners cannot construct grammars incorporating parameter values that are not realised in their L1 because learners only have access to UG via their L1 core grammars. This claim and approach may receive further support from the results of L1 French speakers learning L2 English if we agree with the common view that French is a non-null subject language.¹²⁸ That is to say, because their L1 and L2 share the same value

¹²⁸ French is commonly considered to be a non-null subject language, but it has subject clitics. Subject clitics in French have long been a subject of discussion and debate among linguists (e.g., Kayne, 1975; Auger, 1994; Borer, 1984; Jaeggli, 1982; Zribi-Hertz, 1994). Basically there are two different approaches to account for their syntactic nature: the cliticization approach and the affix approach. Under the cliticization approach, which considers French as a non-null subject language, subject clitics are base-generated in

of the parameter [-null subject], the French learners managed to fix the parameter eventually in response to L2 input.

This view, however, contradicts the second scenario, namely the Parameter Resetting Hypothesis (Schwartz and Sprouse, 1996), which maintains that new parameter values can be reset as IL grammars are fully UG-constrained. Hence, restructuring is predicted to occur at the end-state grammars of L2 speakers, even in cases when the L1 and L2 differ in parameter values. According to this view, based on the observed gradual improvements in the learners' performance at different proficiency levels sharing the same L1, one could argue that they would readjust their L1 value of the null subject parameter to the value appropriate to their L2 with increased proficiency in L2. In other words, because the advanced-level learners have a greater command of the target grammar under investigation compared to their less-proficient upper-intermediate-level peers, who themselves outperform their less proficient lower-intermediate-level counterparts of the same linguistic background, it could be argued that convergence to native-like performance will eventually come. However, this argument seems to contrast with the results obtained in particular from the advanced-level Finnish and Arabic speakers, which indicate that their IL grammars still diverge from the grammar of English native speakers even though they are in the advanced stage of L2 development. However, this contrast can be explained by the fact that it is extremely difficult to determine if an L2 learner is really at the end-state of the developmental process; a learner might be somewhere in the advanced level of proficiency, yet not at the end-state (see Lardiere, 1998a, 1998b; Long, 2003; and White, 2003 for relevant discussion).¹²⁹ To

the canonical subject position within VP, and then undergo movement to their preverbal position. However, under the affix view, which considers French as a null subject language, subject clitics are considered to be agreement prefixes. For reasons of space limitations, I will not comment on any particular position in this thesis. I will go with the common assumption that French is a non-null subject language. However, the interested reader is referred to the references above; or for a review, he/she is referred to Prévost (2009).

¹²⁹ See also the methodological problem discussed in section 4.3.5 in relation to measuring L2 learners' levels of language proficiency. Despite the satisfactory solution proposed to solve this problem, it should be added at this point that the only way to

illustrate this notion, the advanced-level learners' data need to be analysed at the individual level to see if there is some individual variation in performance within the same L1 group. In fact, difference in performance is predicted on an individual basis in the light of the maximum and minimum number of null subject acceptances reported in the subgroups' descriptive analyses presented in subsection 5.2.1.1 above. For example, Table 5.3 showed that the advanced-level Finnish participants' acceptance ranged from zero to three sentences with referential null subjects on an individual basis.¹³⁰ Figure 5.9 precisely illustrates the differences in performance among every individual advanced Finnish L2 learner of English.

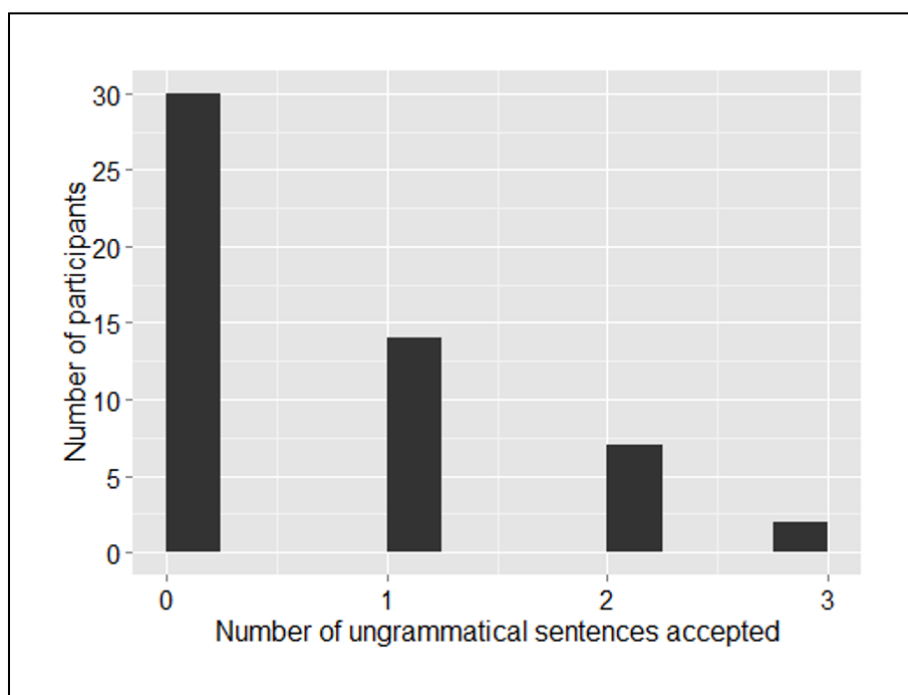


Figure 5-9. Number of sentences with null subjects accepted by every individual Finnish advanced participant.

Figure 5.9 shows that 30 out of 53 Finnish advanced-level participants performed like English natives; they accepted no ungrammatical sentences. The

determine if a learner indeed reached steady-state grammar is by means of longitudinal data. This view is also stated in Lardiere (1998a, 1998b).

¹³⁰ Due to the large number of participants and space limitations, only the descriptive results from Finnish advanced-level participants are presented at the individual level.

other 23 participants accepted sentences with null embedded subjects. While 14 out of those accepted only one sentence with mistakes, seven participants incorrectly accepted two sentences. The other two participants each accepted three sentences with mistakes. Therefore, it could be argued in agreement with the Parameter Resetting Hypothesis that because the majority of the participants converged to native-like usage, the others need more time to reset the parameters as they have not yet reached the end-state.

The discussion so far has shown that these results must be interpreted with caution. Even though the second scenario does not explain the discrepancy between acceptance of null subjects and the absence of null subjects in the learners' production, especially at the advanced level of proficiency, the explanation based on differences in the task demands would account for such discrepancy/inconsistency only at the early stages of the acquisitional process. Therefore, I will further analyse the data before drawing a conclusion about the issue of null subject parameter resetting. Until then, this issue will be left open; I will return to it after discussing how the learners' results diverge across the different syntactic formations under investigation. This will be the focus of the following section of the chapter.

5.3 Results and discussion by grammatical constructions

The primary purpose of this section is to further investigate the data to see if different syntactic structures bring about different performances with regard to null subject production and/or acceptance in L2A. In other words, this section is meant to further investigate the subgroups' results to see whether null subjects are equally spread out across the different syntactic structures under investigation (i.e., complement clauses vs. adverbial clauses). Therefore, the results for each subgroup of learners were broken down further by the item types. This procedure assists an understanding of how learners diverge across the different syntactic formations as far as embedded

subjects are concerned and whether they would converge to English native-like use of subject pronouns in one of these grammatical constructions.¹³¹

This section consists of two subsections. The first subsection deals with the results of translation tasks comparing the performance of misuse of English subject pronouns in adverbial clauses with those in complement clauses within each subgroup; the second deals with the results of the GJ task in the same way as the results on the translation task were analysed and interpreted. Each subsection contains two subsections to first present and then discuss the results in light of some of the theoretical assumptions outlined in Chapter 3, namely L1 transfer and parameter resetting.

5.3.1 The translation task: Subject drop in complement clauses vs. in adverbial clauses

5.3.1.1 The English IL of the French, Finnish and Arabic speakers compared within subgroups

In light of all the descriptive data presented in subsection 5.2.2, which demonstrated that all three groups of learners of all proficiency levels showed high success rates at overtly producing the referential embedded subject pronouns when translating the given sentences into English, it can be predicted that the different L2 syntactic constructions under investigation (complement clause vs. adverbial clauses) have no effect on the learners' performance as far as embedded referential subject pronouns are concerned. The groups' results presented above in subsection 5.2.2 are summarized below in Table 5.13 in terms of mean scores of subject drops.

Table 5-12. Summary table of mean scores of sentences with null subjects produced by French, Finnish, and Arabic L2 learners of English

L1	Lower-intermediate	Upper-intermediate	Advanced
French participants	0.11	0.10	0.00
Finnish participants	0.83	0.18	0.00
Arabic participants	1.75	0.64	0.18

¹³¹ Note the results on the previous subsections have shown that the learners did not manage to converge to native-like if the L1 and L2 differ in parameter values.

The subgroups' low averages use of null referential subjects can be misleading in the sense that it is possible that subject drop might have occurred more frequently in a certain tested syntactic construction but not in another. Therefore, this prediction that the learners' performances do not vary from grammatical construction to grammatical construction cannot be assumed until a comparison within each subgroup is made between the learners' performance in complement clauses and in adverbial clauses with reference to referential subject drop.

After breaking down the results of the translation task's experimental sentences by item types (complement vs. adverbial), the new descriptive statistics of the three lower-intermediate-level subgroups of learners with respect to their use of null referential subjects are presented below in Table 5.14. It shows that both the lower-intermediate French participants and lower-intermediate Finnish participants as subgroups drop subjects at a slightly higher rate on sentences with adverbial embedded clauses (1.1% and 7.5%, respectively) than on those with complement embedded clauses (0.0% and 3.8%, respectively). However, the inferential tests conducted to compare the performance within each subgroup indicated that there were no significant effects for item type. The results arrived at via the Wilcoxon tests were $P\text{-value} = 1.000 > 0.05$ for the French lower-intermediate learners and $P\text{-value} = 0.371 > 0.05$ for the Finnish lower-intermediate learners, which suggests that those subgroups of learners treated missing subjects equally across the different investigated contexts.

Table 5-13. Descriptive statistics by items types (complement vs. adverbial) for sentences with null subjects produced by lower-intermediate subgroups of learners

L1	Items type	Included translation	No. of omission	Max	Min	Mean	Std Dev	Percentage of omission
French	Complement	45	0	0	0	0.00	0.0	0.0
	Adverbial	90	1	1	0	0.11	0.3	1.1
Finnish	Complement	26	1	1	0	0.17	0.4	3.8
	Adverbial	53	4	3	0	0.67	1.2	7.5
Arabic	Complement	75	2	2	0	0.13	0.5	2.7
	Adverbial	149	26	7	0	1.63	2.3	17.4

Unlike their French- and Finnish-speaking counterparts, the Arabic-speaking lower-intermediate participants as a subgroup noticeably dropped subjects at a much

higher rate on sentences with adverbial clauses than on those with complement clauses. As Table 5.14 shows, subject omission occurs in 17.4% of the total included sentences with adverbial clauses, whereas it occurs only in 2.7% for sentences with complement clauses. The result of the Wilcoxon test confirmed that Arabic-speaking learners as a subgroup performed significantly differently in the two types of experimental sentences ($P\text{-value} = 0.005 < 0.05$). The finding that lower-intermediate Arabic learners drop fewer subject pronouns in sentences with complement clauses can be explained by the fact that, in Arabic, a pronominal bound form, which functions as the subject of the complement clause, must be cliticised onto the complementiser *ʔanna* ('that').¹³²

Interestingly, when these lower-intermediate-level results of the Arabic-speaking learners were compared with the results of the English native controls via the Wilcoxon tests, the inferential results show that their performance differed significantly from that of the native English speakers only on sentences with subjectless adverbial clauses ($P\text{-value} = 0.010 < 0.05$).¹³³ The fact that the difference between the lower-intermediate subgroup and the native group did not achieve significance on sentences with subjectless complement clauses ($P\text{-value} = 0.571 > 0.05$) indicates that they performed within the native-like ranges at the lower-intermediate-level on this type of structure as far as null subjects are concerned.

As the results in subsection 5.2.2 indicate, both the French and the Finnish lower-intermediate learners performed native-like with regard to the referential embedded overt subject; it is expected that their performance would still be within native-like ranges even after their results were broken down further by sentence

¹³² For more detail about resumptive pronouns in Arabic, see subsection 2.3.2.

¹³³ Statistically speaking, the results of the native control presented in Tables 5.2 above need not be broken down further by the item types because they did not accept null subjects and therefore did drop them. Thus, where it is required to see how their results diverge from the natives' and when their performance converged to native-like, the learners' results for each item type in this section will be compared directly against the control group's results presented in Table 5.2 for the GJ task to avoid repetition and for space reasons.

structural types. This expectation was confirmed when the inferential tests were applied; all the results obtained by the Wilcoxon tests indicate there were not significant differences between these lower-intermediate participants' performances and the English control with respect to subject dropping in the different structures tested. Consider the results in Table 5.15.

Table 5-14. Inferential comparisons between the lower-intermediate French and Finnish participants and the native controls: subject drop in complement clauses vs. in adverbial clauses

Inter-subgroup comparisons	Clause type	Inferential test	P-value
LI French participants vs. native English controls	Complement	Wilcoxon test	N/A*
	Adverbial	Wilcoxon test	0.450>0.05
LI Finnish participants vs. native English controls	Complement	Wilcoxon test	0.355>0.05
	Adverbial	Wilcoxon test	0.140>0.05

*N/A= not available (the results of both groups were identical: zero number of subjects dropped).

It should be mentioned before leaving this subsection that the translation task results of both the upper-intermediate and advanced subgroups were intentionally not presented here because these two proficiency subgroups of learners, regardless of their linguistic backgrounds, performed either completely native-like or very close to native-like with respect to overt referential subject pronouns in English, as shown above in subsection 5.2.2. This implies that the results of these more proficient learners would not reveal any information to confirm or disconfirm whether L2 syntactic structures would bring about different performance with regard to pronominal subject production in L2A.

5.3.1.2 Discussion of the translation task's results by grammatical structures

The results and discussion in the previous subsection show that the L2 syntactic constructions under investigation have no effect in the performance of both the lower-intermediate French and Finnish L2 learners of English as far as embedded subjects are concerned. Conversely, the lower-intermediate Arabic participants were found to drop subject pronouns differently in the two types of experimental sentences,

dropping subjects at a much higher rate in sentences with adverbial clauses than in those with complement clauses.¹³⁴

Therefore, as both the French and Finnish lower-intermediate learners performed native-like in the two types of experimental sentences and since the performance of the lower-intermediate Arabic participants was attributed to transfer from Arabic, these findings confirm the prediction of hypothesis H3, which stated that different L2 syntactic structures might not bring about different performance on the overt and/or null realisation of the embedded subject pronoun. However, this hypothesis cannot be confirmed on the basis of the findings that emerged from the translation task alone because

- i. the translation task appeared to be quite easy compared to the GJ task, which has proven to be much more difficult for participants regarding embedded subjects (refer to the results and discussion in section 5.2)
- ii. Even though no significant effect for item type (complement vs. adverbial) was found in either the French the Finnish lower-intermediate learners' data, there seems to be a tendency in these groups to drop more subjects in sentences with adverbial clauses than in those with complement clauses. For the sake of illustration, the distribution of the data presented above in Table 5.14 is presented graphically in Figure 5.10 below with respect to the percentage of null referential subject pronoun use.

¹³⁴ Some possible explanations to account for why the L1 Arabic-speaking L2 learners of English dropped more subjects in adverbial clauses than in complement clauses are proposed below, on pages 185-187.

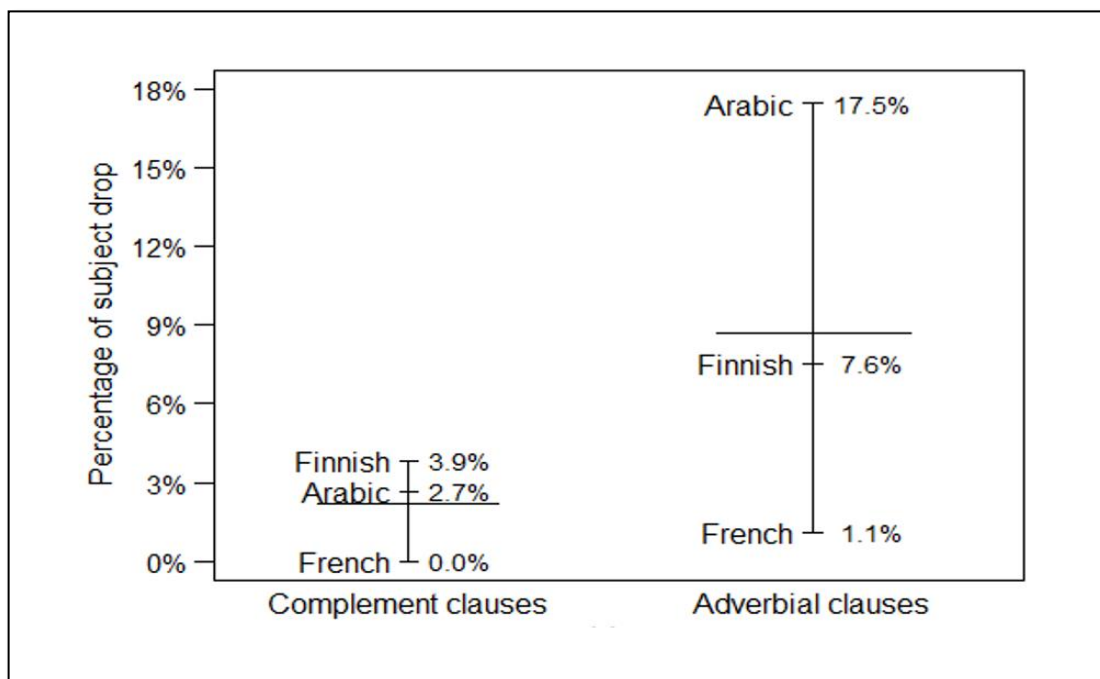


Figure 5-10. Percentages of subject drop in complement clauses and in adverbial clauses by the French, Finnish, and Arabic lower-intermediate learners of English.

Therefore, we need to examine the groups' results obtained via the GJ task in both syntactic contexts to be able to see an even clearer picture before a conclusion can be drawn about the possible effects of L2 sentences structures on learners' performance. This will be the focus of the following subsection.

5.3.2 The Grammaticality Judgment Task: Null Subject Acceptance in Complement Clauses vs. Adverbial Clauses

5.3.2.1 The English IL of the French, Finnish, and Arabic speakers: Compared within subgroups

i. The English IL: French speakers

After breaking down the results of the GJ task by sentences' syntactic constructions (complement vs. adverbial) and based on the descriptive data provided in Table 5.16 below, it seems that acceptance of sentences with null embedded subjects is not equally distributed across different syntactic contexts in the French participants' data. The results show that the learners at all levels of proficiency appear to perform better at rejecting null subjects in the complement clauses than in the adverbial clauses.

Table 5-15. L1 French acceptances by level of sentences with null subjects in embedded complement and adverbial clauses

Item-Type	PL*	Included responses	No. of acceptance	Max	Min	Mean	Std Dev	Acceptance percentage
Null subjects in complement clauses	LI	40	9	4	0	1.00	1.50	22.5
	UI	163	4	1	0	0.13	0.34	2.5
	ADV	161	1	1	0	0.04	0.19	0.6
Null subjects in adverbial clauses	LI	187	103	14	8	11.4	2.46	55.1
	UI	639	89	8	0	2.87	2.01	13.9
	ADV	587	13	5	0	0.48	1.12	2.2

*PL = Proficiency level

Notably, in judging the ungrammatical sentences with missing embedded subjects, the French lower-intermediate subgroup accepted 55.1% of the sentences involving adverbial clauses. In sentences with complement clauses, the same subgroup accepted only 22.5% of the errors. The upper-intermediate subgroup's non-target performance with regard to embedded pronouns was 13.9% in judging ungrammatical adverbial constructions and 2.5% in judging ungrammatical complement structures. In determining the ungrammaticality of sentences involving adverbial clauses, the advanced subgroup accepted 2.2% of the items compared to 0.6% of the accepted items with complement constructions. For the sake of illustration, the French speakers' results presented in Table 5.16 are represented graphically in Figure 5.11 below in terms of acceptance percentages of null subjects across the two tested syntactic constructions.

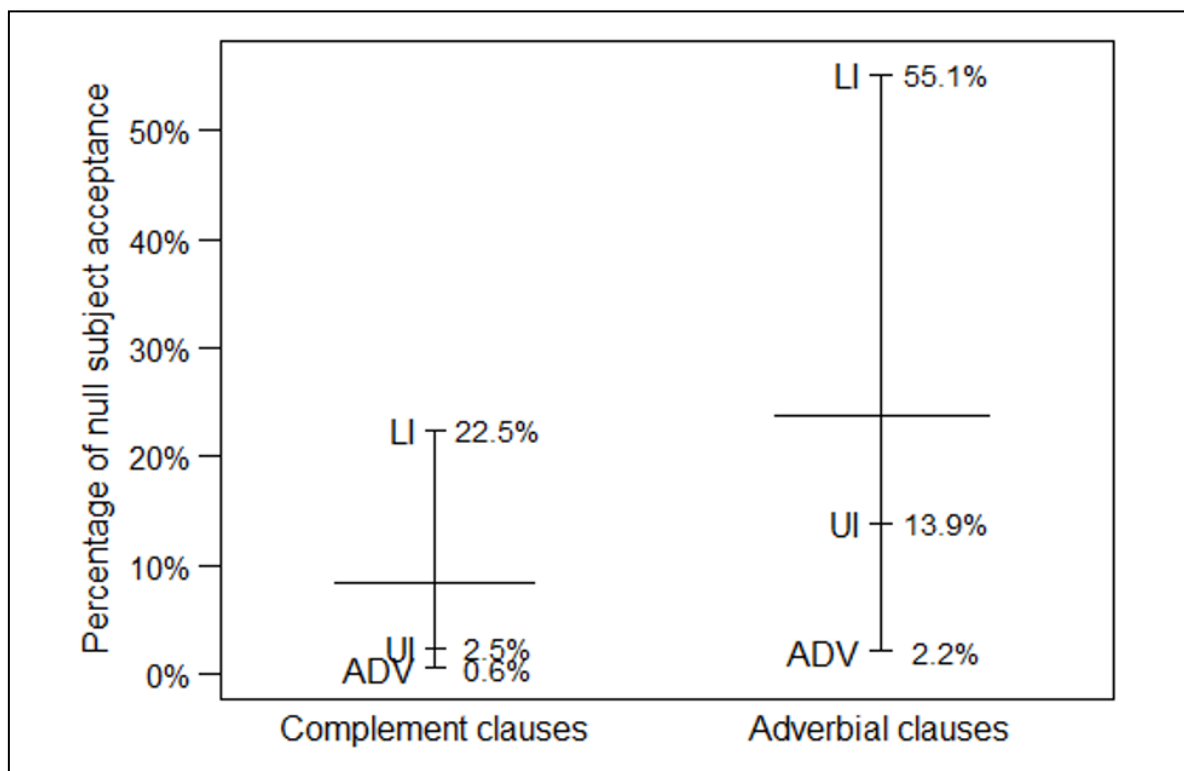


Figure 5-11. L1 French acceptance by level of sentences with null subjects in embedded complement and adverbial clauses.

To investigate the effect of sentence structure on the French-speaking learners' performance, Wilcoxon tests were employed to compare their performance within each subgroup to measure the complement clauses against adverbial clauses with respect to judging ungrammatical items with embedded null subjects. The inferential results all indicate significant effects for sentence types: P-value = 0.009 < 0.05, P-value = 0.000 < 0.05, and P-value = 0.049 < 0.05 for the lower-intermediate learners, upper-intermediate learners, and advanced learners, respectively.

When inter-subgroup comparisons were conducted to compare the results of each proficiency subgroup (Table 5.16) with the results of the English native control group (Table 5.2), the analytical results obtained via the Wilcoxon tests indicate that the differences in the acceptance of complement clauses with null subjects between the L1 French speakers, as individual subgroups of proficiency, and the native English speakers did not reach statistical significance. However, when it comes to acceptance of null subjects in adverbial clauses, only the advanced subgroup did not differ from the native control group under this test. Consider Table 5.17, where highlighted in grey are those L2 subgroups of learners who did not behave like

English native speakers regarding acceptance of null subjects in the different tested grammatical constructions.

Table 5-16. Comparisons between the French participants and the native English controls: acceptance of null subjects in complement vs. in adverbial clauses

Inter-subgroup comparisons	Proficiency level	Inferential test	P-value
L1 French participants vs. native English controls: On complement clauses	LI	Wilcoxon	0.060>0.05
	UI	Wilcoxon	0.339>0.05
	ADV	Wilcoxon	0.663>0.05
L1 French participants vs. native English controls: On adverbial clauses	LI	Wilcoxon	0.001<0.05
	UI	Wilcoxon	0.000<0.05
	ADV	Wilcoxon	0.189>0.05

ii. The English IL: Finnish speakers

The Finnish-speaking learners exhibited similar acquisitional patterns to that of French-speaking learners in that, they, as subgroups of different proficiency levels, accepted noticeably fewer null subject sentences involving complement clauses than they did ungrammatical sentences involving adverbial clauses. As Table 5.18 below and Figure 5.12 (which reports the results graphically via the percentage of acceptance of null referential embedded subjects) show, the lower-intermediate group accepted 38.2% of the ungrammatical sentences involving embedded subjectless adverbial clauses, whereas the scores for the other groups were upper-intermediate 8.2%, and advanced 1.0%. As for the ungrammatical structures containing embedded adverbial clauses, the lower-intermediate accepted 48.3%, the upper-intermediate 12.8%, and advanced 2.7% of the mistakes.

Table 5-17. L1 Finnish acceptances by level to sentences with null subjects in embedded complement and adverbial clauses

Item-Type	PL	Included responses	No. of acceptance	Max	Min	Mean	Std Dev	Acceptance percentage
Null subjects in complement clauses	LI	34	13	4	0	2.17	1.60	38.2
	UI	98	8	4	0	0.47	1.01	8.2
	ADV	313	3	1	0	0.06	0.23	1.0
Null subjects in adverbial clauses	LI	118	57	13	5	9.50	3.67	48.3
	UI	352	45	7	0	2.65	2.18	12.8
	ADV	1159	31	3	0	0.58	0.84	2.7

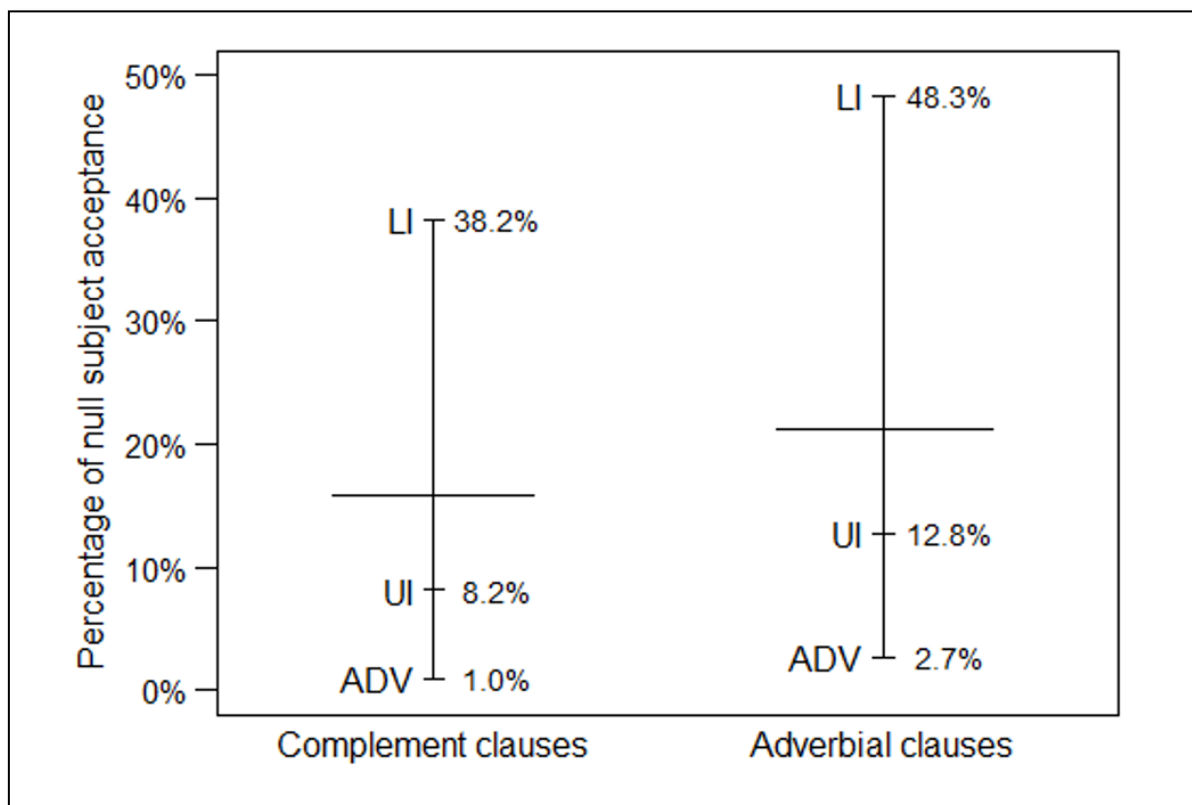


Figure 5-12. L1 Finnish acceptances by level of sentences with null subjects in embedded complement and adverbial clauses.

The inferential tests used for comparing the effects of the item structural types on performance confirmed that there were significant differences within each subgroup of proficiency between the acceptances of embedded null subjects in complement clauses and those in adverbial clauses. The result arrived at via the t-test is $P\text{-value} = 0.016 < 0.05$ for the Finnish lower-intermediate learners; the results obtained by the Wilcoxon tests are $P\text{-value} = 0.001 < 0.05$ and $P\text{-value} = 0.000 < 0.05$ for the Finnish upper-intermediate-level and advanced-level learners, respectively.

When the results in Table 5.18 are statistically compared with the results of the English native control presented in Table 5.2, via the Wilcoxon tests, the results indicate that the L1 Finnish participants of all proficiency levels as individual subgroups differed significantly from native English speakers in the acceptance of adverbial clauses with null subjects. However, when it comes to acceptance of null subjects in complement clauses, only the lower-intermediate-level subgroup differs from the native control group under this test, but not the upper-intermediate and the advanced subgroups. Consider the following table, where highlighted in grey are those Finnish subgroups of participants who did not behave like English native

speakers with regard to acceptance of null subjects in the different tested grammatical constructions.

Table 5-18. Comparisons between the Finnish participants and the native English controls: acceptance of null subjects in complement vs. in adverbial clauses

Inter-subgroup comparisons	Proficiency level	Inferential test	P-value
L1 Finnish participants vs. native English controls: On complement clauses	LI	Wilcoxon	0.005<0.05
	UI	Wilcoxon	0.127>0.05
	ADV	Wilcoxon	0.542>0.05
L1 Finnish participants vs. native English controls: On adverbial clauses	LI	Wilcoxon	0.001<0.05
	UI	Wilcoxon	0.001<0.05
	ADV	Wilcoxon	0.047<0.05

iii. The English IL: Arabic speakers

The same acquisitional trend is noticed when the results obtained by the Arabic-speaking learners are considered. Similar to their French and Finnish counterparts, all three Arabic subgroups performed with higher accuracy rates at rejecting ungrammatical items including embedded complement subjectless clauses than ungrammatical sentences involving adverbial clauses. Table 5.20 compares the incorrect acceptance of null embedded subjects in these two constructions.

Table 5-19. L1 Arabic participants acceptances by level of sentences with null subjects in embedded complement and adverbial clauses

Item-Type	PL	Included responses	No. of acceptance	Max	Min	Mean	Std Dev	Acceptance percentage
Null subjects in complement clauses	LI	87	44	6	0	2.75	2.02	50.7
	UI	77	9	4	0	0.64	1.15	11.7
	ADV	64	1	1	0	0.09	0.30	1.6
Null subjects in adverbial clauses	LI	310	225	20	9	14.1	3.71	72.6
	UI	284	87	13	1	6.21	3.60	30.6
	ADV	235	13	4	0	1.18	1.40	5.5

As Table 5.20 illustrates, when judging ungrammatical items with missing embedded pronominal subjects, the Arabic lower-intermediate subgroup accepted 50.7% of those sentences involving complement clauses. With regard to the sentences involving adverbial clauses, the same subgroup accepted 72.6% of the

mistakes. The upper-intermediate subgroup’s non-target performance with regard to embedded pronouns was 11.7% in judging ungrammatical complement constructions and 30.6% in judging ungrammatical adverbial structures. In judging the ungrammaticality of sentences involving complement clauses, the advanced subgroup accepted 1.6% of the items compared to 5.5% of accepted items with adverbial constructions. For better comparative visualization, these percentages of incorrect acceptance are presented below in Figure 5.13 via Boxplot.

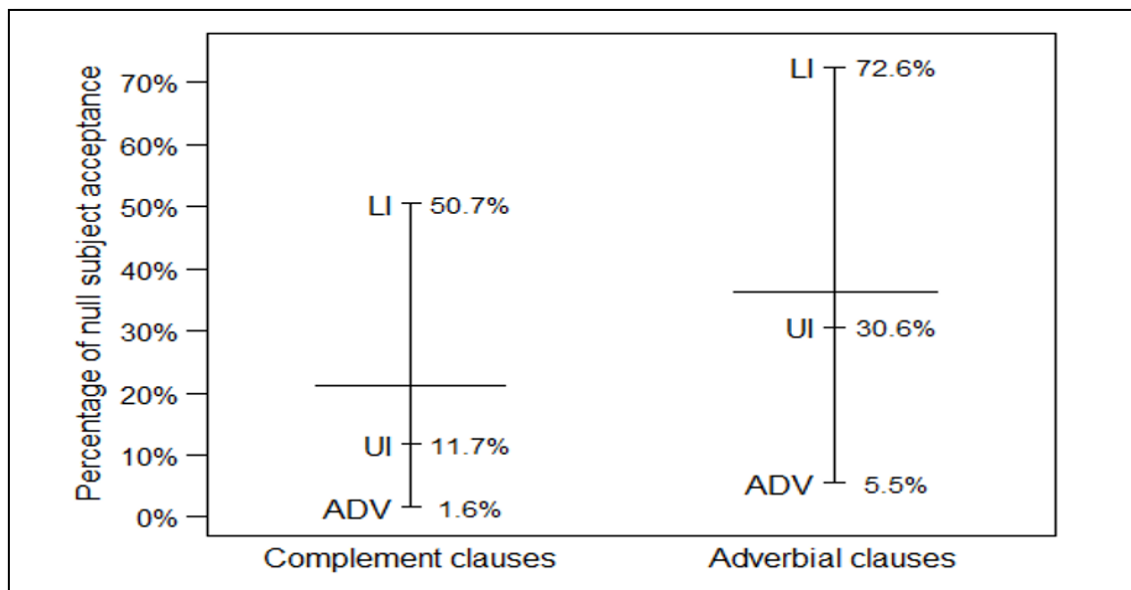


Figure 5-13. L1 Arabic-speaking participants’ acceptances by level of sentences with null subjects in embedded complement and adverbial clauses.

Figure 5.13 suggests statistically significant differences within each subgroup between the learners’ performance on the subjectless complement clauses and their performance on the subjectless adverbial clauses; these differences were later confirmed in the inferential tests. The result arrived at via the t-test is $P\text{-value} = 0.000 < 0.05$ for the Arabic-speaking lower-intermediate learners; the Wilcoxon tests results are $P\text{-value} = 0.001 < 0.05$ for the UL-level subgroup and $P\text{-value} = 0.046 < 0.05$ for the advanced-level subgroup.

As with the Finnish participants, when the results of the Arabic-speaking participants within their subgroup (illustrated in Table 5.20) were compared with the results of the English native control group (Table 5.2), the Arabic-speaking participants at all individual proficiency levels differed significantly from native English speakers in the acceptance of adverbial clauses with null subjects, as indicated by the inferential results arrived at via a series of Wilcoxon tests. However, when it

comes to acceptance of null subjects in complement clauses, only the lower-intermediate-level subgroup differs from the native control group under this test, but not the upper-intermediate and the advanced subgroups of learners. In Table 5.21, highlighted in grey are the subgroups of L2 learners that did not behave like English native speakers with regard to acceptance of null subjects in the different tested grammatical constructions.

Table 5-20. Comparisons between the Arabic-speaking participants and the native English controls: acceptance of null subjects in complement vs. adverbial clauses

Inter-subgroup comparisons	Proficiency level	Inferential test	P-value
L1 Arabic participants vs. native English controls: On complement clauses	LI	Wilcoxon	0.001<0.05
	UI	Wilcoxon	0.089>0.05
	ADV	Wilcoxon	0.494>0.05
L1 Arabic participants vs. native English controls: On adverbial clauses	LI	Wilcoxon	0.000<0.05
	UI	Wilcoxon	0.000<0.05
	ADV	Wilcoxon	0.027 0.05

5.3.2.2 Discussion of the GJ task's results by grammatical structures

The comparison of the performance in the acceptance of null subjects within each subgroup conducted in the previous subsection made it clear that all subgroups, regardless of their L1 backgrounds and their proficiency levels, treated null subjects in the two types of experimental sentences differently (refer to Tables 5.16, 5.18 and 5.20); they accepted far more null subjects in adverbial clauses than in complement clauses (see also the discussion below). These inconsistencies in performance across the two types of experimental sentences disconfirm the prediction of hypothesis H3, which predicted that learners would not be sensitive to the different L2 grammatical structures when judging sentences with null-embedded subjects in English.

I am not aware of any study that has been done to determine if there are structural constraints placed by L2 grammar that could prevent, restrict, and/or allow null subject transfer, to which the present study could be compared; however, the finding that these learners have more difficulties judging certain structures than others may explain the different controversial and inconclusive results produced by L2 studies that use GJ tasks to investigate null subject transfer and parameter

resetting (see Chapter 3). Additionally, no study has been conducted to investigate the frequency of embedded null subjects either in Finnish discourse or in Arabic discourse—it may well be that they occur more frequently in adverbial clauses than complement clauses. This will be left for future research. What is more, the UG theory does not offer any explanation why null subjects would be more frequent in adverbial clauses compared to complement clauses. However, in trying to meet the UG theory's requirement of explanatory adequacy, I will propose some possible explanations from outside the theory to account for why the L2 learners of English had fewer problems in detecting the ungrammaticality of sentences with null embedded referential subjects in complement clauses than in adverbial clauses.¹³⁵ It should be acknowledged, though, before proposing the following explanations that none of these can be assumed until further empirical investigations are conducted to test their credibility to account for such contextual contradictory judgements:

- i. It is possible that null subjects in embedded adverbial clauses were more likely to be accepted than null subjects in complement clauses because the main clauses they are linked to generally provide more discourse contexts compared to the main clauses linked to complement clauses. It follows that the referents of the embedded null subjects in adverbial clauses are more likely to be recoverable from the preceding discourse than the referents of the embedded null subjects in complement clauses. Compare the following examples from the GJ task:
 1. *John will not marry until finds the right woman.
 2. *I told them that feel sick.

It is clear for discourse-pragmatic reasons that the covert subject of the embedded adverbial clause in (1) refers to the same subject of the matrix clause, *John*. However, the silent subject in (2) could refer to different external referent/s from the one in focus (the first-person pronoun *I*) in the preceding discourse, such as *you, we, they, John and Mary*, and so on. A follow-up study

¹³⁵ Explanatory adequacy refers to the requirement that any adequate syntactic theory must explain why grammars have the syntactic properties they do and how acquirers come to acquire such grammars. For more information on this notion, refer to Adger 2003 or Radford 2004.

would be required to test this possibility. It could be tested in a straightforward manner by comparing L2 learners' performance on sentences with one argument plus an adverbial clause with sentences that have two arguments plus an adverbial clause, or comparing L2 learners' performance for sentences that do and sentences that do not have two arguments plus a complement clause.

- ii. Another appealing explanation for this observation could come from the difference between what Davidson (1996) calls pragmatic weight and epistemic parentheticals. Davidson (1996) uses the term pragmatic weight to refer to speakers' use of overt subject pronouns to show that their utterances are "more personally relevant and more vested with emotion (p. 555) ... and [to] increase their "stake" in whatever they are saying" (p. 551). He found that native Spanish speakers used overt pronouns to add pragmatic weight to their speech with verbs of claiming, belief, opinion, and knowledge. On the contrary, other verbs classified as epistemic parentheticals, such as verbs of knowing, seeing, or watching, which "are prone to becoming 'bleached' of their truth-functional content, and which develop more 'abstract' meanings, serving to give information about how the speaker positions him- or herself in relation to their utterance" (Davidson, 1996, p. 557). In his analysis of conversational data, Davidson noticed that native Spanish speakers prefer to use null subject pronouns with epistemic parentheticals.

Surprisingly enough, though this was unintentional, the verbs used in the present study in sentences with complement clauses were similar to the ones in which the Spanish speakers in Davidson's (1996) study used overt pronouns to add pragmatic weight to their utterances; the verbs used in the sentences with adverbial clauses were also similar to the ones with which the Spanish speakers used null subjects. This is illustrated in the following examples from the experimental sentences.¹³⁶

1. a. *She claims that fell in the water.
b. *They believe that got good marks.

¹³⁶ For more information about the verbs used within the embedded clauses in the GJ task, refer to Appendix 1a.

2. a. *He watched the movie until fell asleep.
- b. *They must not talk while watch TV.

This explanation may be partly illustrated by the notion of psychological focus of attention. A pronoun is in focus “if the attention of both speech participants can be assumed to be focused on it” (Gundel, 1999, p. 294). It could therefore be argued that overt pronouns are used to add pragmatic weight with verbs of claiming, belief, opinion, and knowledge because such verbs are likely to bring a referent into focus.

If this reasoning explains why these L2 learners accepted far more adverbial clauses with null subjects than complement clauses with null subjects, this would raise questions worth investigating: Do L2 learners transfer their L1 discourse-pragmatics roles (i.e., pragmatic weight and epistemic parenthetical)? Do learners acquire the discourse-pragmatics roles that govern the distribution of null and overt subjects in L2? If yes, which ones are acquired first and which are learned later?

- iii. Haegeman (2010) noticed that certain types of adverbial clauses in English are highly independent of the main clause, specifically temporal versus conditional clauses. From this perspective, it could be argued that learners were more tolerant of null subjects in adverbial clauses because such clauses are more independent than the complement clauses, again for discourse and pragmatic reasons (see the argument in (i) above).

It is far from clear at this stage which of these alternatives (or other unstated ones) offers a more adequate explanation for the existing discrepancy in the learners’ judgements of sentences with null subjects in complement clauses or adverbial clauses. I will leave this question and other related questions for future research.

We can now shift our attention to the parameter resetting issue to continue the discussion left open in subsection 5.2.3.2 above. However, before doing so, we need to again consider the results presented in the present subsection that focus on whether the IL grammars converged in all respects on the target grammar. A distinction is observed in the acceptance of null subjects in two types of the

experimental sentences; the learners of all proficiency levels, regardless of their L1 backgrounds, accepted significantly fewer null subjects in complement clauses than in adverbial clauses. The French participants performed completely native-like with respect to rejecting null subjects in complement clauses despite the fact that they accepted them in adverbial clauses at both the lower-intermediate- and upper-intermediate-levels; however, they converged on the target grammar in all respects at the advanced-level. In contrast, both the Finnish- and Arabic-speaking participants accepted complement clauses with null subjects only at the lower-intermediate-level; they both managed to converge on native-like usage of overt subjects in this particular grammatical construction at the upper-intermediate-level. On the contrary, these two L1 groups of learners continued to perform non-native-like in their judgment of null subjects in adverbial clauses in that they continued to accept null subjects in such clauses even at the advanced stage of L2 acquisition.

Such a contrast is not predicted under the Parameter Resetting Hypothesis; if the null subject parameter can be reset to a value appropriate to the L2, why do Finnish and Arabic participants establish the requirement for overt subjects in English complement clauses early on, but persistently continue to accept null subjects in adverbial clauses even at the advanced stage of L2 development? To put it another way, if a new parametric value can be acquired, it should be expected that null subject acceptance would disappear from the learners' judgments in both grammatical structures at the same developmental stage, or at least they would disappear from that particular grammatical structure persistently accepted at the advanced stage of the acquisitional process.

It seems, therefore, that the findings from the GJ task are compatible, in principle, with the No Parameter Resetting Hypothesis. The discrepancy between the learners' acceptance of null subjects in complement clauses and adverbial clauses suggests that null subjects cannot be fully eliminated from the learners' advanced IL grammars. This implies that the various groups of Finnish and the Arabic speakers have failed to reset their L1 value of the null subject parameter to the L2 value of the parameter. However, this raises the question of how one can explain the observed gradual improvement in learners' performance in relation to their proficiency if the L1 value of the null subject parameter cannot be reset. In other

words, what strategies did some advanced L1 Finnish and Arabic learners of English use in order to make their performance completely native-like compared with their advanced peers, who themselves managed to bring their performance much closer to native-like performance than their less proficient counterparts (lower-intermediate- and upper-intermediate-level learners)? The answer to this question is not straightforward. However, it could be argued that the learners are becoming increasingly better at applying a conscious learning strategy to choose the overt option for pronouns consistently.¹³⁷ In principle they could become so good at applying this strategy that they would be indistinguishable from the native controls in this experiment. The prediction is that even the participants who performed at a native-speaker's level in our experiment would still occasionally lapse into using a null subject or accepting a null subject if they were retested again using sentences with different types of grammatical structures. This argument is compatible with the

¹³⁷ L2 researchers (Krashen (1981), Schwartz (1993), De Keyser (2003), White (2003), Paradis (2004), Ullman (2005) among many others) argue for two separate types of knowledge: explicit and implicit knowledge. Explicit knowledge forms based on conscious information/learning (knowledge grows by means of explicit instruction), whereas implicit knowledge grows in response to unconscious information (knowledge triggered by exposure to linguistic input). These researchers argue that explicit learnt knowledge, because it is assumed to be under conscious control, is not available to learners all the time; "it can only be used in circumstances ... where sufficient time is available and sufficient attention is devoted to the form of the language ...; [in other words, it is used] to monitor what the unconscious system produced and potentially to modify it" (Towell, 2013, pp. 121-132). Therefore, it could be argued since the findings are compatible with the No Parameter Resetting Hypothesis that the learners in the present study depended on their conscious learnt knowledge rather than their unconscious knowledge in order to bring their performance much closer to native-like with time and increased proficiency in L2A. Otherwise, a clear cut-off point would be expected in which the advanced learners would perform at native level if [+prodrop] is not part of the learners' implicit competence of English.

Note L2 generative researchers (e.g. Hawkins and Towell, 1991; Hawkins, 2001 and White, 2003) argue in favour of learning that takes place implicitly triggered by speech input, as it is thought to lead to language mastery - convergence on the target grammar (for more detailed discussion about the distinction between the process of learning and the process of acquisition, refer to Towell, 2013).

hypothesis that parameters cannot be reset, but it also does not exclude the possibility that they could be. This kind of argument is also compatible with the Full Transfer/Full Access Hypothesis (Schwartz and Sprouse, 1994, 1996) and the Modulated Structure Building approach (Hawkins, 2001a). These hypotheses suggest that steady-state IL grammar might converge on the target language grammar or diverge from that grammar depending on the learner's L1 properties and TL input. However, such a divergent grammar is fully UG-constrained; it might include properties from the L1 grammar, the L2 grammar, or any other grammar (for more detail, see Chapter 3).¹³⁸

However, the results presented in this chapter reflect only the overall developmental trend for the groups of learners; group results are not necessarily representative of the totality of the individual learners in the group. Therefore, the end-group conclusion on the whole cannot be individualised. This is because language acquisition is an individual construct, a long process that is affected at the individual level by numerous variables internal and external to the learner. Hence, it is possible that L2 learners may end up with different UG-constrained IL competences, even learners with the same L1 who have learned the same L2.

To illustrate the problem of generalising observation, let us reconsider as an example the individual results of the advanced Finnish learners of English to see if the general conclusions drawn based on the group results, which shows that these learners as a subgroup could not reset the value of the null subject parameter, can be applied to every individual learner in the subgroup. Figure 5.14 below demonstrates their individual performances with respect to accepting null subject in adverbial clauses.¹³⁹

¹³⁸ To better understand this theory, consider the relationships among initial and end-states in SLA under the No Transfer/No Access Hypothesis of Epstein, Flynn, and Martohardjono (1996), which maintains that the IL grammar at the end-state must be native-like. Consider also the research concerns with the issue of fossilisation in SLA.

¹³⁹ The learners' individual performances in complement clauses were not analysed here because the Finnish participants converged on native-like usage of overt subjects in this

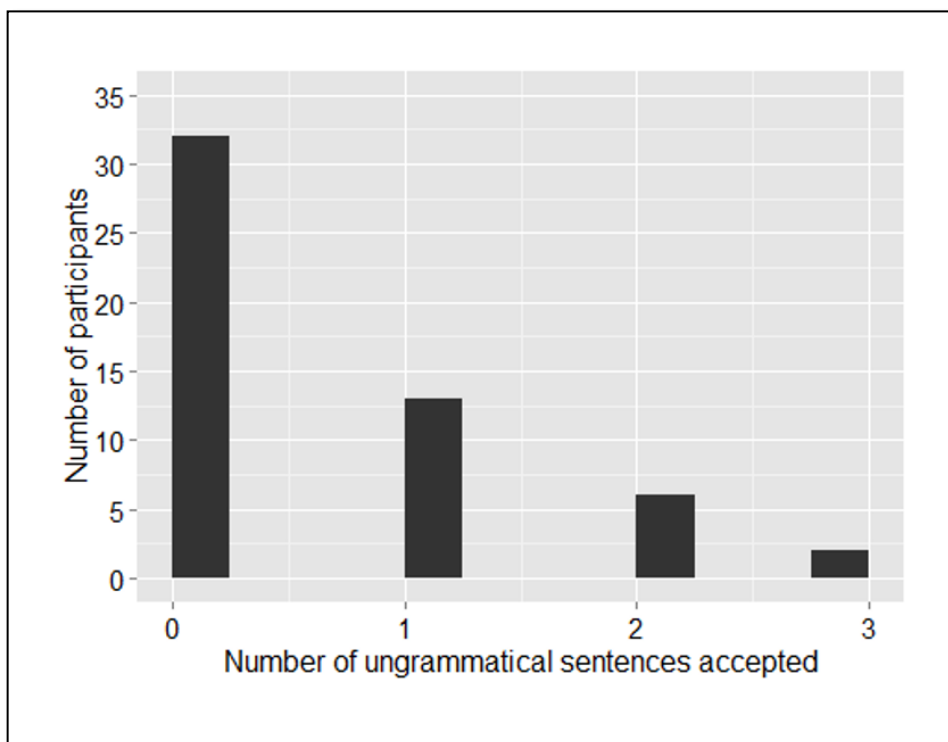


Figure 5-14. Number of embedded adverbial clauses with null subjects accepted by Finnish advanced-level learners of English.

Figure 5.14 shows that 32 out of 53 Finnish advanced-level participants accepted none of the embedded adverbial clauses with null subjects. The other 21 participants accepted such ungrammatical clauses in varying degrees: 13 accepted only one sentence with mistakes, six participants incorrectly accepted two sentences, and the remaining 2 participants accepted three sentences with mistakes. These data illustrate that null subject acceptances are not equally spread out across the judgments of every individual learner. The results reveal that while some participants demonstrated native-like performance, others performed non-native-like in their use of overt pronouns, although some were closer to native-like usage compared with the rest.

grammatical structure early on after passing the lower-intermediate stage of development (see discussion in subsection 5.3.2). Therefore, their individual performances on this particular syntactic structure at such an advanced developmental stage do not illustrate the problem of results generalisation.

Therefore, in order not to be misguided by the group observations that suggested the null subject parameter setting cannot be reset to a value appropriate to the L2, and in order to control the problem of variation in performances among the speakers of the same L1 learning the same L2, it is better to argue that individual learners may attain native-like performance.

As we are driven by our desire to find universalities in human cognition, the above discussion on the notion of generalisation leads us to another problem. Similar to the observation that the group-study's finding cannot be individualised, the group-study's finding cannot be generalised beyond the grammatical domain investigated in that study – namely, the null subject parameter. For example, we cannot assume that all other L1 parameters values (i.e., the Word-Order Parameter which is also termed as the Head Position Parameter) cannot be reset at the group level based on our observation that the null subject parameter cannot be reset at the group level.¹⁴⁰ This sort of argument can be supported on the one hand by Hawkins, Towell, and Bazergui's argument (1993) that "certain aspects of UG instantiated in the L1 are highly resistant to revision in L2 learning while others are readily revisable by L2 learners" (p. 220). On the other hand, this view can be supported indirectly by the findings of the present study that L2 learners have more difficulty judging certain structures than others, which suggests that the grammar of an L2 may contain constraints that could prevent, restrict, and/or allow the acceptance of null subjects in L2A. This finding implies that the possibility of parameter resetting not only depends on the similarities and differences between L1 and the properties of L2 and TL input but also on the other constraints and properties of the TL grammar.

¹⁴⁰ The Word-Order Parameter refers to the relative positioning of heads with respect to their complements in a given language – either be head-first or head-last (For more detail, see Radford 2004). It should be mentioned that it has been found that adult L2 learners can reset their L1 value of this parameter to the value appropriate to the L2 easily (refer to Hawkins 2001a).

5.4 Referential null subjects licensing in SLA

The results in the previous section have shown that there are structural, developmental, and situational/contextual (realised as task-type) constraints on when, where, and to what extent null subject parameter settings are transferrable. Having tried to provide plausible explanations for the existence of these constraints earlier in this chapter, this section seeks to investigate the mechanisms by which null subjects are licensed in the learner's IL grammar. To examine the issue in depth, learners' performances were investigated and discussed under the principles of different theories put forward to account for the null subject phenomenon, namely the formal grammar approach, which focuses on the potential connection between verbal agreement and null subjects (Chomsky, 1981); performance-based approaches, especially those that focus on processing capacity limitations (Bloom, 1990; Valian, 1991; Valian and Eisenberg, 1996); and the informational context approach, which focuses on discourse-pragmatic factors regarding null subjects, including the preceding discourse and the informational value of the subject and the situational context, including the speaker/hearer (Margaza and Bel, 2006; Montrul, 2004; Sorace, 2004; Tsimpli, Sorace, Heycock, and Filiaci, 2004).

In view of the theoretical link between rich verbal agreement morphology and null subject licensing in some languages including Arabic and Finnish (discussed in Chapter 2), it seems reasonable to begin by investigating the learners' knowledge of agreement morphology to see if they transferred their L1 agreement system to English to license null subjects. Therefore, we need to examine the developmental path(s) of verbal agreement inflection and compare it/them with the developmental path(s) of null subjects. If Agreement (AGR) is the licenser of null subjects, one would expect that a smaller number of sentences without subject-verb agreement would be accepted compared with sentences with null subjects, as this means they have acquired English agreement/projected English AGR. In other words, if there is a relationship between these two variables, we would expect an inverse (negative) relationship where a high rate of null subject acceptance correlates with a low rate of missing subject-verb agreement acceptance.

However, before examining the knowledge of projected English AGR, comparing the results from subject-verb agreement with the results from null

subjects, and computing the correlation between these two variables using the Pearson product-moment coefficient, we need to first redo the group statistics with respect to acceptance and production of null subjects (summarised in Figures 5.1 and 5.4 above), including exclusively 3rd-person singular null subject pronouns, because agreement features in English are only overtly marked for 3rd-person singular by the present indicative –s.¹⁴¹ There were 17 ungrammatical sentences of this type in the GJ task and 6 sentences in the translation task, after excluding those items with pronominal connection to 1st and 2nd person and those with pronominal reference to 3rd-person plural from the analysis of sentences with null subjects accepted and produced by the participants (see Chapter 4).

Figure 5.15 below presents the acceptance distribution in percentage on this type of items with 3rd-person singular null subject pronouns. For the purpose of quantitative comparisons, this figure also presents the distribution of acceptance of ungrammatical sentences with missing subject-verb agreement.¹⁴²

¹⁴¹ First- and second-person singular pronouns agree with verbs “*to be*” (am, is, and are); however, such cases were not part of the investigation in the present study.

¹⁴² Due to space limitations, and because AGR was not the licenser of null subjects in the learners’ performance (see discussion below), the detailed descriptive statistics for these two variables will not be presented here. However, relevant tables illustrating and further comparing the learners’ performance within each subgroup in the two elicitation tasks can be found in Appendix 11.

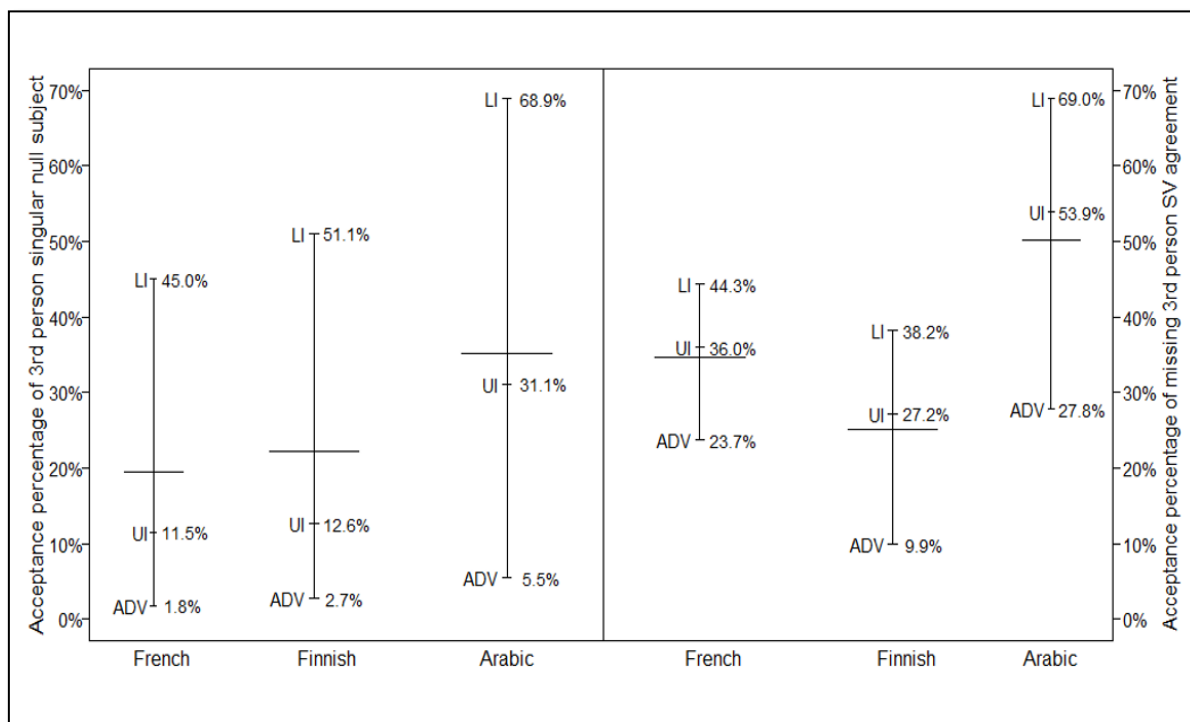


Figure 5-15. Participants' IL systems: Percentage of acceptance of items with third-person singular null subjects and items with missing subject-verb agreement.

Figure 5.15 clearly shows, based on the comparison of the performance within the different subgroups, that only the French and Finnish lower-intermediate subgroup participants accepted fewer sentences with missing subject-verb agreement than they did sentences with null subjects. All seven other subgroups of learners exhibited the reverse pattern; they accepted more sentences involving missing subject-verb agreement than sentences with null subjects. This finding suggests that these L2 learners had more problems in general in detecting the ungrammatical sentences with missing subject-verb agreement than with null subjects, which in turn means that null subjects are still produced before the mastery of agreement morphology. Thus, one can argue that there is no relationship between null subjects and subject-verb agreement in SLA.¹⁴³ This sort of argument also receives support

¹⁴³ One of the examiners suggested investigating the correlation between individual subjects' performance in identifying null-subject errors, and their performance in identifying agreement errors; he mentions that even if we don't see this pattern within individuals, it doesn't mean that there isn't a relationship in terms of the diachronic development of the language. I have not followed this recommendation because it is hard to look at each participant's individual results due to the large number of participants involved in this study (see section 4.3.1 and section 5.1). As for diachronic development,

from the behaviour of the French and Finnish participants. It would be expected that, if AGR were the licenser of null subjects in their judgments, both the upper-intermediate- and advanced-level French and Finnish learners would have exhibited the same patterns of performance exhibited by the lower-intermediate-level French and Finnish learners until they fully mastered the function of 's inflectional morphology; they would be expected to accept fewer sentences with missing subject-verb agreement than they did sentences with null subjects, not the reverse, which they did. Therefore, the performance of the French and Finnish lower-intermediate participants, which at the beginning suggested that AGR played a role in licensing null subjects, can be explained by the argument that these lower-intermediate-level learners appeared unaware of the necessity of the 's inflectional morphology in such structures in English. This sort of argument is supported by the very high percentages of sentences with missing subject-verb agreement (44.3% and 38.2%) and with null subjects (45% and 51.1%) that the French and Finnish lower-intermediate learners accepted (see Figure 5.15). Therefore, AGR was not the licenser of the null subjects in the lower-intermediate French and Finnish learners' data.

However, we cannot arrive at the conclusion that there is no association between verbal agreement and null subject acceptance only on the basis of the fact that numerical—but not statistical—differences existed within each subgroup of learners between the acceptance in their data of items involving the two variables. Therefore, Pearson product–moment correlations were computed between these two variables—null subject acceptance and missing subject-verb agreement acceptance—to investigate the relationship and to measure the strength and direction of the association between them within the performance of each subgroup of learners. Table 5.22 displays the results of the correlations analyses.

I acknowledge that he is right. However, it is a difficult question to say anything about at this stage because measuring a learner's performance at one point in time using techniques such as the GJ task or a translation task provides only an incomplete picture of the processes of language development. Therefore, the most satisfactory method to investigate how ILs develop over time (i.e., to investigate the possible existence of such a relationship between null subjects and subject-verb agreement in terms of the diachronic development of the language) is to conduct a longitudinal study.

Table 5-21. Correlations between third-person singular null subjects and missing subject-verb agreements

L1	Subgroup	P-value	Correlation Coefficient
French	LI	0.227>0.05	0.447
	UI	0.015<0.05	0.433
	ADV	0.930>0.05	-0.018
Finnish	LI	0.579>0.05	-0.289
	UI	0.047<0.05	0.488
	ADV	0.004<0.05	0.387
Arabic	LI	0.653>0.05	0.122
	UI	0.032<0.05	0.574
	ADV	0.127>0.05	0.489

Following the line of reasoning that only an inverse relationship reflects a developmental connection between null subjects and subject-verb agreement, the results presented in Table 5.22 reveal no significant negative correlation between these variables, which confirms the initial descriptive finding that AGR is not the licenser of null subjects in L2A. This observation is consistent with the findings of previous L2A studies (Clahsen and Hong, 1995; Davies, 1996; Meisel, 1991; White, 1985, 1986) as well as the findings emerging from L1 acquisition studies (Ingham, 1998; Radford, 1990; Sano and Hyams, 1994; Valian, Hoeffner, and Aubry, 1996).

Note that the positive significant correlations, in which the percentage of null subject acceptance is proportional to the percentage of missing subject-verb agreement acceptance,¹⁴⁴ found between the two phenomena within the performance of the four subgroups of learners (highlighted in grey in the table above) might be explained by the fact that learners at the upper-intermediate and advanced stages of acquisition usually have more stable rules and can therefore make more accurate judgments. However, this does not mean that these learners discover the impoverished agreement system of English and are on their way to block *pro*; this is not only because the Finnish and the Arabic participants as subgroups continue to accept null subjects even at the advanced stage of acquisition but also because the

¹⁴⁴ As the acceptance of null subjects decreases, the acceptance of missing subject-verb agreement decreases as well.

results indicate that the learners' mastery of agreement features is quite limited. This observation is consistent, in principle, with the general predictions of the hypotheses put forward to account for the failure or persistent problems of uninterpretable features in L2A (e.g., The Missing Surface Inflection Hypothesis [Prévost and White, 2000], The Representational Deficit Hypothesis [Hawkins, 2005; Hawkins and Hattori, 2006] and The Feature Reassembly Hypothesis [Lardiere, 2008, 2009]).¹⁴⁵ Evidence supporting the claim that uninterpretable features are difficult to acquire in L2 acquisition can be found in a statistical comparison within each subgroup the overall acceptance of 3rd-person singular null subjects with the overall acceptance of missing subject-verb agreement. Consider Table 5.23:

Table 5-22. Inferential comparisons between acceptance of third-person singular null subjects and acceptance of missing subject-verb agreements

L1	Proficiency level	Inferential test	P-value
French	LI	Paired t-test	0.013<0.05
	UI	Paired Wilcoxon test	0.002<0.05
	ADV	Paired Wilcoxon test	0.000<0.05
Finnish	LI	Paired t-test	0.010<0.05
	UI	Paired t-test	0.217>0.05
	ADV	Paired Wilcoxon test	0.005<0.05
Arabic	LI	Paired t-test	0.000<0.05
	UI	Paired t-test	0.922>0.05
	ADV	Paired Wilcoxon test	0.041<0.05

As table 5.23 shows, the results reveal that only the upper-intermediate Finnish- and Arabic-speaking participants show no significant differences between the acceptance of these two variables; the seven other subgroups groups performed significantly better in rejecting the sentences with third-person singular null subjects than those with missing subject-verb agreement.¹⁴⁶ This confirms the finding that the acquisition of overt subject and the acquisition of agreement features in L2 English are independent of each other. This finding is further confirmed by the learners' performance in the translation task, as illustrated in Figure 5.16. Note that all

¹⁴⁵ Refer to Chapter 3 for more details about these theories.

¹⁴⁶ Consider the results of these tests in the light of the quantitative differences between the acceptances of these two variables, as illustrated by Figure 5.15 above.

subgroups of learners performed better in the use of overt subjects than in the use of agreement morphology in that they produced far fewer sentences with embedded null subjects than they did sentences with missing subject-verb agreement.

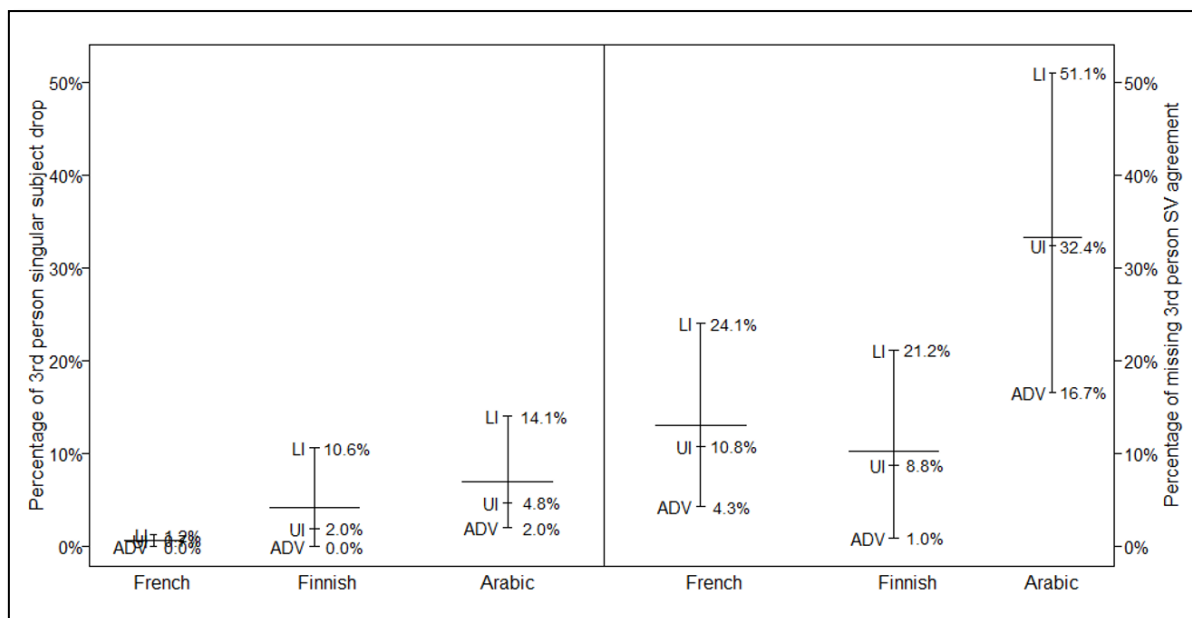


Figure 5-16. Learners' performance on the translation task: Percentages of third-person subject drops vs. percentages of missing subject-verb agreement.

The discussion thus leads us to the conclusion that the syntactic approach cannot offer an explanation accounting for the existence of null subjects in the learners' IL grammars of L2 English. This conclusion is also supported by the empirical data presented and discussed in the present chapter in sections 5.2 and 5.3 above. From a syntactic perspective, an approach based on the richness of AGR cannot explain why the participants accept more null subjects in adverbial clauses than in complement clauses in accordance with the results discussed in section 5.3 or why they generally accepted more null subjects with local antecedents than with non-local antecedents in accordance with the results found in subsection 5.2.1.3. These results also indicate that null subjects in L2A are unlikely to be explained by a performance deficit account; if adult L2A were constrained by processing capacity limitations, these learners (or at least the advanced ones) would not be expected to differ in their treatment of null subjects depending on the type of the clause involved (complement vs adverbial) or the position of the antecedent involved (local vs non-local). This is because under the performance deficit account, it is expected that learners would accept null subjects in longer complex utterances more often than in

shorter ones (see section 2.4); however, since both complement and adverbial clauses are syntactically complex sentences (see section 4.3.2.1), learners are expected to treat null subjects in these different syntactic structures similarly if there is a performance-deficit explanation for this phenomenon.

In fact, these results show that only a discourse-pragmatic approach can offer acceptable explanations for such inconsistent performances. In subsection 5.2.1.3 and section 5.3, these questions (i.e., why the participants accepted more null subjects with local antecedents than with non-local antecedents and why they accepted more null subjects in adverbial clauses than in complement clauses) were discussed, mainly based on an informational context approach that focused on the discourse-pragmatic factors relating to missing subjects. Such a sensitivity to discourse-pragmatic factors in subject realisation has indeed been reported in some recent L2 developmental literature (see Chapter 3; cf. also Margaza and Bel, 2006; Montrul, 2004; Quesada and Blackwell, 2009; Sorace, 2004; Tsimpli, Sorace, Heycock, and Filiaci, 2004).

The superiority of this account to all other competence-based or performance-based accounts is that it can predict when and where a subject is more likely to be dropped or accepted as null in language acquisition. A null subject is more likely to occur when its referent is unambiguously inferable from the discourse context; if its referent is not or is less recoverable from the linguistic context, this argument is more likely to be realised overtly. Thus, argument omission versus overt expression depends on the referent's discourse status, which can be defined in terms of a range of discourse and pragmatic notions including person, focus of attention, topic, topic-shift, contrast, and animacy among many other notions (Ariel, 1990, 1996; Chafe, 1996; Davidson, 1996; Frascarelli, 2007; Gundel, 1999; Gundel, Hedberg, and Zacharski, 1993; Kempson, 1996; Lambrecht, 1994). Greenfield and Smith (1976) capture the complex relationship between the referent's discourse status and argument realization in their Principle of Informativeness, which states that informative subjects whose referents are not highly inferable either through the linguistic or the extralinguistic contexts are much more likely to be realised overtly than uninformative subjects whose referents are associated with old information that are well established in the discourse or non-linguistic context. To illustrate this, consider as an example

the role played by person features in argument realization. It has been documented (cf. Bhat, 2007; Bianchi, 2006; Chafe, 1994, 1996; Dimitriadis, 1995) that first- and second-person referents are always unambiguously identified by the speaker or the hearer, whereas third-person referents can be ambiguously identified. Ambiguity arises, for example, when two or more referents can be potential antecedents. Therefore, in null subject languages third-person subjects are more likely to be realised overtly in order to avoid ambiguities at a discourse level compared to first- and second-person subjects, which tend to be realized covertly. In line with this argument, Allen (2000) and Hughes and Allen (2006) found that Inuktitut-speaking children and English-speaking children are much more likely to realise third-person subjects overtly than either first- or second-person subjects.

However, as opposed to this argument, and related to the findings from child L1A, all subgroups of learners who participated in the present study, regardless of their linguistic backgrounds and proficiency levels, accepted significantly more embedded 3rd-person null subjects than 1st- and 2nd-person null subjects in L2 English.¹⁴⁷ This finding also disconfirms the prediction of hypothesis H1, which predicted that the Finnish participants would accept/drop more null subjects in the 1st- and 2nd-person contexts compared to those in the 3rd-person contexts. Figure 5.17 below summarises the learners' results as subgroups on the GJ task in terms of percentage of acceptance of null referential embedded subjects for third-person pronouns versus first- and second-person pronouns.¹⁴⁸

¹⁴⁷ Due to the fact that null subjects were generally rare events in the participants' production (primarily limited to lower-intermediate participants), the translation results have not been analysed for the effects of person features in argument realization on the learners' performance.

¹⁴⁸ Again, due to space limitations, the detailed descriptive statistics for the two variables under the discussion will not be presented here; however, the relevant tables comparing learners' performances within each subgroup can be found in Appendix 12.

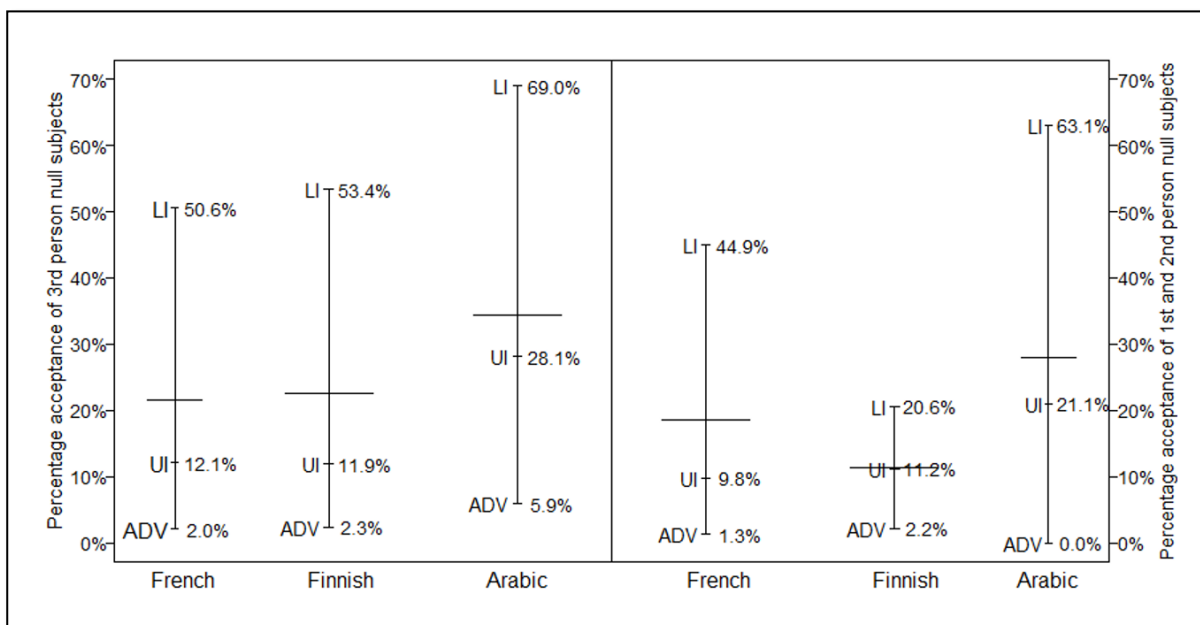


Figure 5-17. Intra-subgroup comparison of non-target performance in the acceptance of null subjects with third-person referents and with first- or second-person referents in the GJ task.

Table 5-23. Statistical intra-subgroup comparisons between acceptance of third-person null subjects and acceptance of first- and second-person null subjects

L1	Proficiency Level	Inferential Test	P-value
French	LI	Paired t-test	0.000<0.05
	UI	Paired Wilcoxon test	0.000<0.05
	ADV	Paired Wilcoxon test	0.034<0.05
Finnish	LI	Paired t-test	0.000<0.05
	UI	Paired Wilcoxon test	0.007<0.05
	ADV	Paired Wilcoxon test	0.008<0.05
Arabic	LI	Paired Wilcoxon test	0.000<0.05
	UI	Paired Wilcoxon test	0.002<0.05
	ADV	Paired Wilcoxon test	0.034<0.05

Table 5.24 above presents the inferential results of the comparison of the performance within subgroups, revealing that all subgroups of learners were significantly better at rejecting ungrammatical sentences with embedded 1st- or 2nd-person null subjects than ungrammatical structures with embedded 3rd-person null subjects controlled by higher arguments.¹⁴⁹

This unexpected finding raises the following three questions:

- i. If learners relied on discourse and pragmatic factors to license null subjects, why did they not obey the predicted person feature roles in argument realization? The answer to this question leads to the second related question below.
- ii. Why is a higher null subject acceptance rate found in 3rd-person embedded contexts in learners' performance?
- iii. Why do L1 acquirers exhibit different developmental patterns from L2 learners with respect to the variables under investigation, namely omitting a higher number of subjects in 1st- or 2nd-person contexts as opposed to L2 learners who accept higher numbers of null subjects in the 3rd-person contexts?

¹⁴⁹ However, when the results presented in Figure 5.17 above are submitted to further statistical analysis using the generalised linear mixed effects models (GLMM) - an interaction analysis applied to further investigate whether the performance on the variables under investigation (acceptance of 3rd-person null subjects and 1st- or 2nd-person null subjects) differ depending on the learners' L1 and proficiency levels - the inferential results reveal that only the lower-intermediate Finnish-speaking participants show a significant relationship between their L1 and their acceptance of 3rd-person null subjects and 1st- or 2nd-person null subjects. So, future research is necessary to investigate the issue in depth before a conclusion can be drawn based on these results. Consider the following table:

Table 5-24. The relationship between L1 and acceptance of 3rd-person null subjects and 1st- or 2nd-person null subjects

L1	Proficiency Level	P-value
French	LI	0.364
	UI	0.674
	ADV	0.801
Finnish	LI	0.000
	UI	0.831
	ADV	0.967
Arabic	LI	0.927
	UI	0.901
	ADV	0.348

As for the first question, it is well documented that grammatical person is only one of several discourse and pragmatic factors that interact with each other and with the syntax to legislate when subjects can be null (see Ariel, 1990, 1996; Chafe, 1996; Davidson, 1996; Frascarelli, 2007; Gundel, 1999; Gundel, Hedberg, and Zacharski, 1993; Kempson, 1996; Lambrecht, 1994). To illustrate how the interaction between the different discourse and pragmatic factors affects argument realization, let us consider in addition to the grammatical person another factor known as *focus* (Gundel et al, 1993). A referent is brought into focus “if the attention of both speech participants can be assumed to be focused on it because of its salience at a given point in the discourse” (Gundel, 1999, p. 294). Gundel et al. (1993) argued that null pronouns are mainly associated cross-linguistically with the “in focus” status. They also mentioned that “subjects and direct objects of matrix sentences are highly likely to bring a referent into focus” (p. 279). Therefore, it could be argued, because the present study focuses only on null subjects in embedded contexts, that the level of ambiguity may rise as the identification of 3rd-person referents is minimized because the 3rd-person null subjects of the embedded clauses maintain their discourse referents from the immediate matrix clauses (refer to the experimental sentences in Appendix 1).

Having seen that the subjects and/or the direct objects of matrix sentences used in the experimental tasks helped to disambiguate the 3rd-person referents, this leads to the discussion to the second interrelated question as to why all the participants accepted significantly more null subjects in the 3rd-person embedded contexts than in the 1st- and 2nd-person embedded contexts, contra our prediction. One possibility for such behaviour can be partly explained in terms of Charney’s person-role hypothesis (1980), which predicts that the pronoun referring to the speaker [+speaker] would be mastered first due to its frequent use, followed by the second-person pronoun [+hearer] due to its direct relevance to the conversational situation; the third-person pronoun would appear later due to the fact that it lacks direct personal features connected to participation in the dialogue [—speaker —hearer]. Hence, in line with this argument, even though this study does not look strictly at the acquisition of subject pronouns, the mastery of subject pronouns may follow a natural

developmental sequence in adult L2A, a developmental observation that can explain why far more 3rd-person null subject sentences were accepted.¹⁵⁰

However, it seems that the person-role hypothesis does not sufficiently account for why the rate of null subject acceptance is substantially higher for 3rd-person than for 1st- and 2nd-person pronouns. It fails to explain why even learners at the advanced stage of proficiency persistently accepted significantly more null subjects in third-person contexts; learners at this stage of acquisition are expected to have acquired all the personal pronoun forms. Another problem with this hypothesis emerges when the relevant L1 acquisition research literature is considered. It has been documented cross-linguistically by many first language acquisition researchers (Bretherton, McNew, and Beeghly-Smith, 1981; Clark, 1978; Guasti, 2002; Rom and Dgani, 1985; Valian, 1991), that children are much more likely to realise 3rd-person subjects overtly than either 1st- or 2nd-person subjects (see Allen, 2000; Hughes and Allen, 2006). Accordingly, it could be concluded that the developmental sequence of subject pronouns offers no solution as to why subject drop rates vary from one personal pronoun to another. This question will be left for future research. However, two points might help explain this behaviour. First, some languages, such as the Lak language and Khasi language, have systems of personal pronouns that only consist of 1st- and 2nd-person pronouns; the 3rd-person pronouns belong to the system of demonstratives (cf. Bhat, 2007). Second, in some other languages, such as Old Norwegian and Old Icelandic, only 3rd-person can be null (cf. Kinn, 2013; Walkden, 2014)

At this point, we shift our attention to the third question: Why do L1 acquirers exhibit different developmental patterns from L2 learners with respect to the variables under investigation, namely omitting higher number of subjects in first- or second-person contexts as opposed to L2 learners who accept higher numbers of null subjects in third-person contexts?

¹⁵⁰ I am not aware of any study that has been done to test this issue in L2A – the developmental sequence of subject pronouns.

Such different acquisitional paths can be explained by the idea that adult L2 learners and child L1 acquirers use null subjects for different reasons in non-null subject languages such as English. While L2 learners rely on discourse licensing of null subjects, children allow null subjects for performance-deficit reasons (see Chapter 2 for more detail about this performance-based account). Evidence that supports this claim comes from empirical evidence that null subjects in child grammar have different distribution from null subjects in adult L2 grammar. It has been observed that children's null subjects are only attested in simple clauses but not in embedded clauses, whereas most null subjects in L2 learners' performance occur in embedded clauses, where a null subject gets its discourse referent from the subject or object of the matrix clause.¹⁵¹ The difference in the range of null subjects, mainly 1st and 2nd persons for children versus mainly 3rd person for L2 learning adults, could be another effect of the different reasons for using null subjects. Indeed, further research is needed to examine the issue and other questions that can be raised in the light of the discussion above regarding the effect of discourse-pragmatics on argument realization in L2: How do L1-specific discourse-pragmatic constraints affect learners' subject realization in L2A? Are there language-specific pragmatic conditions that could govern the distribution of overt and null subjects?¹⁵² If so, can L2 learners of a null subject language acquire the language-specific pragmatic conditions that govern the distribution of overt and null subjects? What does it mean in relation to ultimate attainment if a learner becomes sensitive to the complex discourse-pragmatic factors that require sophisticated communicative ability?

¹⁵¹ See Chapters 2 and 3 for more details about the distribution of null subjects in child grammars and adult IL grammars, respectively.

¹⁵² For a detailed discussion about this point, refer to Cole (2010).

Chapter 6. Conclusion

6.1 Overview of the Chapter

This is the last chapter of the present study. A summary of the empirical experiment, including its goals and its major findings, are presented in section 6.2. Some theoretical and empirical limitations in this study will be addressed in section 6.3. The last section includes some suggestions for future research.

6.2 Summary of the Study

This study was designed to determine whether the L1 null subject parameter value transfers in L2A and whether its value can be reset. For these purposes, it investigates the acquisition of obligatory overt subject pronouns in English by three groups of learners whose L1 belongs to three distinct languages: French (a non-null subject language), Finnish (a partial null subject language) and Arabic (a consistent null subject language). The participants in each group were divided into three subgroups—lower-intermediate, upper-intermediate, and ADV—on the basis of their scores on the proficiency test in order to examine how the investigated L2 grammar changes over time at the different developmental stages in relation to the learners' different native languages. All learners involved in the study were asked to complete a translation task and a grammaticality judgment task. Several novel findings that emerged from the present empirical study are summarized herein.

1. The results of the GJ task showed that all groups of learners—French, Finnish, and Arabic—accepted null subjects in English at the early stages of L2 development despite the fact that their L1s are different in parameter setting. Despite the fact that there were gradual improvements with increased proficiency in the L2 in participants' abilities to recognize the ungrammaticality of such sentences, both the Arabic and the Finnish participants continued to perform non-native-like; only the French participants managed to converge to English native-

like, by hypothesis, as a direct consequence of having an L1 that does not allow null subjects. These findings are consistent with the Modulated Structure Building approach (Hawkins, 2001a) which predicts an initial state bare VP in L2A and at the same time allows for influence of the L1 later at a stage of development when functional projections are posited by the learner. In other words, this approach predicts that the pro-drop parameter setting cannot be represented or transferred at the initial stages of L2A, but can be later transfer of information encoded in the relevant functional projection; in other words, the L1 setting is transferred at the later stages of L2A – could be at the intermediate stage.

2. The grammatical intuition results also indicated the learners differentiated in their acceptance of sentences with null pronouns depending on the position of their referential antecedents. Whereas the French- and the Finnish-speaking learners accepted more null subjects with local antecedents, the Arabic-speaking participants accepted more null subjects with non-local antecedents. Given how pro-drop works in Arabic, which allows external antecedents, as opposed to Finnish, where pro-drop antecedents have to be local and preferably no further away than the preceding clause, the difference between the Arabic and the Finnish participants' performance is expected in agreement with the Modulated Structure Building approach (Hawkins, 2001) which allows for influence of the L1 during the subsequent process of structure building. These differences in performance support the distinction between types of pro-drop discussed in Chapter 2 and also provide evidence for Holmberg's argument (2010a) that there are several null subject parameters.
3. The results from the translation task suggest that the learners from different linguistic backgrounds varied with regards to the production of referential embedded subject pronouns in English – a finding which suggests that L1 parameter setting transfers in L2A. The French participants behaved completely like the native English speakers from early on, by hypothesis, as a direct consequence of having an L1 that does not allow subject omission. Like their French peers, the Finnish participants performed within the native-like ranges from at the lower-intermediate-level, but nevertheless there were marginal differences in their abilities as proficiency subgroups to appropriately translate sentences, unlike their French counterparts, whose abilities to perform the task

consistently were constant from the lower-intermediate-level onwards. This result indicates that the Finnish speakers show a tendency to omit referential subjects more than their French counterparts, although not at a statistically significant level. Such a tendency could have been tested if the L2 initial-state IL grammar were under investigation. Alternatively, such a tendency was statistically confirmed when the results of the Finnish lower-intermediate subgroup were compared with the Arabic lower-intermediate subgroup. Although there was no statistical difference between their performances, there were statistical differences between the performance of the Arabic and the French participants.

4. The contrast emerged when the results of the GJ task were compared to the results of the production task provided evidence that all participants as proficiency subgroups, regardless of their L1 backgrounds, did not perform consistently across different task types, in that their performances varied from task to task. Even though their performances generally became better as proficiency increased, the GJ task seemed to be much more difficult for the L2 learners of English compared to the translation task, which appeared to be quite easy as far as embedded subjects are concerned. The learners persistently accepted referential null subjects in the GJ task beyond the stage of L2 development when they had established the requirement for overt subjects in their production.

5. The results from the grammatical intuitions data showed evidence that different syntactic structures in the L2 bring about different performances on the overt and/or null realisation of the embedded subject pronouns in L2A of English. Learners of all proficiency levels, regardless of their L1 backgrounds, treated null subjects in the two types of experimental sentences differently; they accepted significantly fewer null subjects in complement clauses than in adverbial clauses. The French participants performed completely native-like with respect to rejecting null subjects in complement clauses despite the fact that they accepted them in adverbial clauses at both the lower-intermediate- and upper-intermediate-levels; however, they converged on the target grammar

in all respects at the advanced-level. In contrast, both the Finnish- and Arabic-speaking participants accepted complement clauses with null subjects only at the lower-intermediate-level; they both managed to converge on native-like usage of overt subjects in this particular grammatical construction from the upper-intermediate-level. However, these two L1 groups continued to perform non-native-like in their judgment of null subjects in adverbial clauses; they persistently continued to accept null subjects in such clauses even at the advanced stage of L2 acquisition.

6. The groups-findings, which indicated that L2 learners' performances vary from task to task and from structure to structure (points 4 and 5 above) suggested that null subject parameter setting cannot be reset to a value appropriate to the L2 (Hawkins and Chan, 1997; Smith and Tsimpli, 1995).
7. The results revealed that the grammatical category person plays a role in argument realization in L2A; all the learners, regardless of their linguistic backgrounds and their proficiency levels, were found to differentiate in their treatment of embedded null subjects, depending on their grammatical persons referents (first-, second-, and third-person referents). They unexpectedly accepted significantly more null subjects in the third-person embedded contexts than in the first- and second-person embedded contexts.
8. These findings, which indicate that there are structural and situational or contextual constraints on when and where pronominal subjects can be null, suggest that L2 learners rely on discourse licensing of null subjects. A null subject is more likely to occur when its referent is unambiguously inferable from the discourse context; if its referent is less accessible from the linguistic context, this argument is more likely to be realised overtly. That is to say, argument omission versus overt expression in L2 depends on the referent's discourse status, which can be defined in terms of a range of discourse and pragmatic notions.

I will conclude this section with a very general observation that emerged from this thesis: there are structural, developmental, and situational/contextual (realised as task-type) constraints on when, where, and to what extent pronominal subjects can be null in L2A.

6.3 Limitations of the Study

I need to acknowledge a few limitations of this empirical study. These are grouped into two categories: methodological limitations and theoretical limitations.

1. Methodological limitations

- i. Language acquisition is an individual construct; it is a long process that is affected by numerous variables that are internal and external to the learner. This implies that a learner's performance measured at one point in time using techniques such as the GJ task or a translation task only provides an incomplete picture of the processes of language development. The existence of variability observed in the same individual's performance along with the observed variations in the present study among L2 learners of the same L1 who were grouped under the same level of L2 performance could support this sort of argument. Therefore, perhaps the most satisfactory method to investigate how ILs change over time is to conduct a longitudinal study.¹⁵³
- ii. French is commonly considered to be a non-null subject language, but it has subject clitics.¹⁵⁴ This L1 syntactic feature might have affected the performance of French participants. Therefore, it would have been desirable to get a more accurate picture of the nature of null subjects in L2A by involving another group of learners in the present study whose native language is a non-pro-drop language that does not have subject clitics, such as Swedish.
- iii. The methodology chapter asserted that the knowledge of a second or third

¹⁵³ See Ortega and Byrnes (2008) or Ortega and Ibarra-Shea (2005) for a more detailed argument for using longitudinal data in SLA.

¹⁵⁴ Subject clitics in French has been a subject of debate among linguists (e.g., Auger, 1994; Borer, 1984; Jaeggli, 1982; Kayne, 1975; Zribi-Hertz, 1994). Basically, there are two different approaches to account for their syntactic nature: the cliticisation approach and the affix approach. Under the affix view, subject clitics are considered as agreement prefixes. In that case, there would be a null subject co-occurring with the clitic.

foreign language was a factor that was not completely controlled. Despite the fact that there are few linguistic studies investigating third language (L3) initial state and the subsequent development states, there is disagreement among researchers when there are two syntactic options for a learner to choose/transfer from (L1 and L2) when acquiring an L3. Some researchers concluded that third language acquisition (L3A) is simply another case of L2A; other researchers argue that both L1 and L2 syntactic properties are transferred into L3; the rest of the researchers claim that only L2 syntactic properties are transferred (Flynn, Foley and Vinnitskaya, 2004; Jaensch, 2008; Leung, 2006, 2007; Rothman and Cabrelli Amaro, 2010; Williams and Hammarberg, 1998).

Although this factor might have affected the results to some degree, it was impossible to completely control, particularly in the case of the Finnish participants; in Finland, in addition to Finnish, the language of the majority of the Finnish population, Swedish is also spoken as a national language. The end result of the interaction among the two linguistic communities in some cases leads to bilingualism; as a consequence, some of the Finnish participants acquired Swedish in addition to Finnish natively. However, because Swedish is a non-null subject language like the investigated language, English, all Finnish participants who acquired Swedish as another native language were excluded from the analysis. However, those participants who learned Swedish or any other language(s) after age seven as an L2 were included in the analysis (refer to subsection 4.3.1). Therefore, for future research, we could obtain a more accurate picture about the nature of null subjects in L2A if researchers investigated the knowledge of pronominal subjects in the English of speakers of another partial null subject language such as Brazilian Portuguese, Marathi, or Hebrew.¹⁵⁵

¹⁵⁵ For more details about these partial null subject languages, see Holmberg, Nayudu and Sheehan (2009) and Vainikka and Levy (1999).

2. Theoretical limitations

.Although parameter theory offers a rich theoretical basis that helps L2 linguists to successfully describe and explain various observations about SLA, it fails to explain here why the learners persistently accepted referential null subjects in adverbial clauses beyond the stage of L2 development when they have established the requirement for overt subjects in complement clauses.¹⁵⁶ This finding was not predicted by this theory; in fact, a clear cut-off point would be expected in which the advanced learners would perform at native-level in all structures.

6.4 Suggestions for Further Research

The results of the present study have shown that there are structural constraints on null subject parameter settings transfer in SLA. There is no doubt that, in order to clearly understand the nature of transfer, future research is required to explore in depth how, when, and to what extent L2 syntactic structures condition null subject transfer. In other words, to gain further insight into the nature of missing subjects in IL, it is important to investigate why different L2 syntactic structures bring about different performances on the overt and/or null realisation of subject pronouns in SLA.

The results also raise several questions that require further investigation: Why do the learners persistently accept referential null subjects in adverbial clauses beyond the stage of L2 development when they have established the requirement for overt subjects in complement clauses? Why do the learners accept significantly more null subjects in third-person embedded contexts than in first- and second-person contexts? To what extent does the performance of L2 learners whose L1 is a non-pro-drop language that does not have subject clitics, such as Swedish, learning a non-pro-drop language differ from L2 learners whose L1 is a non-pro-drop language that has subject clitics, such as French, acquiring the same non-pro-drop language? How do first language-specific discourse-pragmatic constraints affect learners' subject realization in L2A? Can L2 learners of a null subject language acquire the language-

¹⁵⁶ This finding is not predicted by any other syntactic theory as far as I am aware.

specific pragmatic conditions that govern the distribution of overt and null subjects?
With reference to ultimate attainment, what does it mean if a learner becomes sensitive to such complex discourse-pragmatic factors that require sophisticated communicative ability?

Appendices

Appendix 1: Tests instruments

Appendix 1.a. The GJ task and the translation task
(French version)

Study investigating knowledge of English

1. Translation Task

Instructions

Here are some sentences in your own language. Please translate **all** of them into English. After each sentence there are some suggestions of words you could use in your translation. If you do not understand a word, please look at the vocabulary list (it has words with their meanings in French). Please follow these models:

1. Mary a rencontré John hier.

meet	Yesterday
------	-----------

Model Answer: Mary met John yesterday

2. Ils aiment jouer dans le parc.

like	Play	Park
------	------	------

Model Answer: They like playing in the park

Enquête sur la maîtrise de l'anglais

1. Exercice de traduction

Instructions

Voici quelques phrases rédigées dans votre langue. Veuillez les traduire toutes en anglais. Après chaque phrase se trouvent des suggestions de mots que vous pouvez utiliser dans votre traduction. Si vous ne comprenez pas un mot, veuillez consulter la liste de vocabulaire (la signification des mots y est donnée en français). Veuillez suivre ces modèles:

1. Mary a rencontré John hier.

meet	yesterday
------	-----------

Exemple à suivre: Mary met John yesterday.

2. Ils aiment jouer dans le parc.

like	play	Park
------	------	------

Exemple à suivre: They like playing in the park

Please do **not** change any sentence after you have translated it. I am interested in your first attempt at translating these sentences. You will have 15 minutes on this exercise; please try to complete all of the sentences.

If you have no questions, please start. Have fun!

* * * * *

The translation exercise starts from the next page.

Je vous prie de **ne pas modifier** une phrase après l'avoir traduite. C'est votre premier essai de traduction de ces phrases qui m'intéresse. Vous disposerez de 15 minutes pour ce test, veuillez essayer de traduire toutes les phrases.

Si vous n'avez pas de questions, veuillez commencer. Amusez-vous bien !

* * * * *

L'exercice de traduction commence à la page suivante.

1. Le verre se brise quand il tombe au sol.

glass	break	when	fall	floor
-------	-------	------	------	-------

.....

2. J'étais très fatigué(e) quand je suis arrivé(e) à la maison.

feel	tired	when	get	home
------	-------	------	-----	------

.....

3. Jane dit qu'elle l'aime.

say	love
-----	------

.....

4. Ma mère s'assoit dans sa chaise quand elle se sent fatiguée.

mother	sit	chair	whenever	feel	tired
--------	-----	-------	----------	------	-------

.....

5. Elle a perdu son livre parce qu'elle l'a oublié dans le train.

Lose	Book	When	leave	train
------	------	------	-------	-------

.....

6. Elle dit qu'elle veut acheter une nouvelle voiture.

Say	want	buy	car
-----	------	-----	-----

.....

7. Il m'appellera quand il arrive.

call	When	arrive
------	------	--------

.....

8. Ils ne quitteront pas le bureau jusqu'à ce qu'ils rencontrent le directeur.

Leave	Office	Until	meet	manager
-------	--------	-------	------	---------

.....

9. John prétend qu'il parle bien anglais.

claim	Speak	English
-------	-------	---------

.....

10. John s'est cassé la jambe quand il a sauté de l'arbre.

break	leg	when	jump	tree
-------	-----	------	------	------

.....

11. Les garçons étaient contents quand ils ont gagné le prix.

boy	happy	when	win	prize
-----	-------	------	-----	-------

.....

12. Marie a aimé le chat quand elle l'a vu.

like	cat	when	see
------	-----	------	-----

.....

13. J'ai dit que j'ai trouvé mon stylo.

say	Find	pen
-----	------	-----

.....

14. Vous pensez que vous avez oublié votre livre dans la classe.

think	leave	book	class
-------	-------	------	-------

.....

15. Tu ne pourras pas regarder de film à la télé, jusqu'à ce que tu as fini de manger

watch	movie	Until	finish	eat
-------	-------	-------	--------	-----

.....

2. Grammatical and Ungrammatical English sentences

Instructions

Here are some sentences. Some are written correctly in English and others contain an error. Please read the sentences carefully and then tick one of the options:

- CLEARLY CORRECT English sentence
- CLEARLY INCORRECT English sentence
- POSSIBLY INCORRECT English sentence
- I DON'T KNOW

Here is an example:

(i) Mr Smith's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Please tick: CLEARLY CORRECT if you are sure that the sentence is GRAMMATICAL (written correctly/ has no error) in English.

2. Phrases anglaises grammaticalement correctes ou non

Instructions

Voici quelques phrases dont certaines sont correctement écrites en anglais, tandis que d'autres contiennent une erreur. Veuillez lire attentivement les phrases puis cocher une des options:

- CLEARLY CORRECT (**MANIFESTEMENT CORRECTE**)
- CLEARLY INCORRECT (**MANIFESTEMENT INCORRECTE**)
- POSSIBLY INCORRECT (**POSSIBLEMENT INCORRECTE**)
- I DON'T KNOW (**JE NE SAIS PAS**)

Voici un exemple:

(i) Mr Smith's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Veillez cocher: CLEARLY CORRECT (MANIFESTEMENT CORRECTE) si vous avez la certitude que la phrase est GRAMMATICALEMENT CORRECTE (rédigée correctement/sans fautes).

Please tick: CLEARLY INCORRECT if you are sure that the sentence is UNGRAMMATICAL (has an error).

Please tick: POSSIBLY INCORRECT if you think there is an error but you are **not certain** about it.

Please tick: I DON'T KNOW if you have **no** idea about the answer.

If you tick **INCORRECT** or **POSSIBLY INCORRECT**, please **draw a line** under the part of the sentence that you believe is wrong. Then, please correct the mistake by rewriting only the wrong part of the sentence. Look at the example below: If you think that the sentence is **incorrect** because of the missing article in 'gave him gift', you would draw a line and add an 'a' as follows:

(ii) Mr Smith's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input checked="" type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: gave him a gift

Veillez cocher : CLEARLY INCORRECT (MANIFESTEMENT INCORRECTE) si vous avez la certitude que la phrase est GRAMMATICALEMENT INCORRECTE (contient une erreur).

Veillez cocher : POSSIBLY INCORRECT (POSSIBLEMENT INCORRECTE) si vous pensez que la phrase contient peut-être une erreur, mais vous n'en êtes **pas certain(e)**.

Veillez cocher : I DON'T KNOW (JE NE SAIS PAS) si vous n'avez **aucune** idée de la réponse.

Si vous cochez **INCORRECT** (INCORRECTE) ou **POSSIBLY INCORRECT** (POSSIBLEMENT INCORRECTE), veuillez **souligner** la partie de la phrase que vous estimez fautive. Ensuite, veuillez corriger l'erreur en réécrivant seulement la partie fautive de la phrase. Consultez l'exemple ci-dessous : si vous pensez que la phrase est incorrecte à cause de l'article manquant dans « gave him gift », vous soulignez cette partie et ajoutez « a », comme suit :

(ii) Mr Smith's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input checked="" type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Devrait être: gave him a gift

Important Notes

1. Remember: If you think the sentence could be **incorrect**, choose POSSIBLY INCORRECT and change the sentence as in example (ii).
2. You have 40 minutes for this test; please make sure you finish all the items.
3. Remember, this test is for research purposes, so you do not need to worry about your score.

Here are three examples to try before you start the test:

(iii) I need to borrow my sister's French-English dictionary.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

(iv) Bill wonders where is Mary going shopping.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

Remarques importantes

1. N'oubliez pas : si vous pensez que la phrase est peut-être incorrecte, cochez POSSIBLY INCORRECT (POSSIBLEMENT INCORRECTE) et modifiez la phrase à la manière de l'exemple (ii).
2. Vous disposez de 40 min pour ce test ; veillez à le terminer entièrement.
3. N'oubliez pas que ce test est destiné à une étude scientifique et que vous n'avez pas à vous inquiéter des résultats.

Voici trois exemples pour vous exercer avant de commencer le test :

(iii) I need to borrow my sister's French-English dictionary.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Devrait être: _____

(iv) Bill wonders where is Mary going shopping.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Devrait être: _____

(v) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

(See the answers below)

Answers

(iii) I need to borrow my sister's French-English dictionary.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

(iii) Bill wonders where is Mary going shopping.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: Mary is

(v) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect.

Devrait être: _____

(Voir les réponses ci-dessous)

Réponses

(iii) I need to borrow my sister's French-English dictionary.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

Devrait être: _____

(iii) Bill wonders where is Mary going shopping.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

Devrait être: Mary is

(vi) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: hopes **the** dictionary.

Do you have questions before you start this test? If you have, please ask the researcher right now. Otherwise, you are now ready to begin. Have fun!

The exercise starts from the next page.

(vi) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

Devrait être: hopes **the** dictionary.

Avez-vous des questions avant de commencer l'exercice ? Si oui, veuillez les poser à l'enquêteur immédiatement. Sinon, vous pouvez commencer. Amusez-vous bien !

L'exercice commence à la page suivante

1. The man drank tea while he read the newspaper.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

2. He drives whenever goes to work.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

3. Mary knows that her brother plays on a football team.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

4. You want to study mathematics at university when leave school.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

5. Anne says that John sleep in class.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

6. Bill says that the doctor told to listen to music.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

7. John claims that George will follow to the farm.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

8. Bill complains that yesterday the teacher stopped from playing football.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

9. John told me that found his money.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

10. The woman said that she arrived in time.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

11. John will not marry until finds the right woman.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

12. The man says that the glass broke when his son played with it.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

13. Linda explained that the library open late at night.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

14. We prefer to swim when we go on holiday in the summer.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

15. The baby often cries when hears loud noise.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

16. We will not leave the house until see her.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

17. You had nearly finished your book when left it on the train.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

18. Jack mentioned that he loves watching TV.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

19. John believed that the university sometimes allow students to study abroad.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

20. Ann was in a very difficult situation when lost her job.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

21. Jane hopes that her father come back.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

22. The rabbit jumped when escaped.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

23. Bill claimed that the professor always give too many low marks.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

24. I know that the car stopped before John filled with petrol.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

25. You can drive while you listen to music.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

26. The cup broke when Adam dropped on the ground.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

27. John believed her when told him the truth.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

28. My son likes the gift when I gave to him.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

29. The manager can see you when finishes his work.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

30. Ann claimed that the cafeteria offer the best food in town.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

31. I had to visit Jane before I left for the holidays.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

32. John said that the postman frequently delivers letters to the wrong house

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

33. Jane thinks that Peter started French classes.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

34. He watched the movie until fell asleep.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

35. Mary always cleans her room before she goes to school.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

36. The children played football until they left.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

37. They must not talk while watch TV.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

38. The boys were happy when won the prize.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

39. The children think that she went to work.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

40. The birds died when tried to escape.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

41. The cat chased the mouse before it ate.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

42. I will not leave his office until pays me the money.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

43. Ann feels that her father hates.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

44. John helped Mary when lost her job.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

45. The girl says that she saw the film.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

46. The vase broke when fell.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

47. Mary said that the swimming pool always close too early.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

48. Paul claims that the man tried to steal money from him.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

49. They slept when got home.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

50. John thought that the cinema often show films in the afternoon.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

51. I cook whenever feel hungry.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

52. The car stopped before hit the child.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

53. My father knows that we watch too much television.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

54. I am always worried that I will miss the train whenever I have to catch.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

55. She will phone us from the airport when arrives.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

56. Sara claims that Jane hit with a book.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

57. Their father saw Susan before returned from school.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

58. John told me that recently he found mathematics quite interesting.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

59. Bill says that the party ended when the police stopped.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

60. John feels that his mother love him.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

61. She claims that fell in the water.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

62. Sara told me that her aunt often reads the newspaper after dinner.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

63. Helen told me that she wants to go to New York

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

64. I told my children that feel sick.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

65. We were unhappy when we came last in the race.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

66. Ann told me that her mother often drink tea after dinner.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

67. The student told me that finished his homework.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

68. You said that left your wife.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

69. Jane says that this morning the teacher allowed to sing.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

70. They believe that got good marks.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

Thank you

Appendix 1.b. The GJ task and the translation task
(Finnish version)

Study investigating knowledge of English

3. The Translation Task

Instructions

Here are some sentences in your own language. Please translate **all** of them into English. After each sentence there are some words that you may use to improve your new English sentences. If you do not understand a word, please look at the vocabulary list (it has words with their meanings in Finnish). Please follow these models:

1. Mary tapasi Johnin eilen.

meet	yesterday
------	-----------

Model Answer: Mary met John yesterday

2. He pitävät puistossa leikkimisestä.

like	play	park
------	------	------

Model Answer: They like playing in the park

Tutkimus englannin kielen tuntemuksesta

1. Käännöstehtävä

Ohjeet

Tässä on muutamia lauseita omalla kielelläsi. Käännä ne kaikki englanniksi. Jokaisen lauseen jälkeen on joitakin sanoja, joita voi käyttää parantaaksesi uusia englanninkielisiä lauseitasi. Jos et ymmärrä jotain sanaa, voit katsoa sanastoa (siellä on sanoja ja niiden merkitys suomeksi). Noudata näitä malleja:

1. Mary tapasi Johnin eilen.

meet	Yesterday
------	-----------

Mallivastaus: Mary met John yesterday

2. He pitävät puistossa leikkimisestä.

like	Play	park
------	------	------

Mallivastaus: They like playing in the park

Please do **not** change any sentence after you have translated it. I am interested in your first attempt at translating these sentences. You will have 15 minutes on this test; please try to complete all of the sentences.

If you have no questions, please start. Have fun!

* * * * *

The translation test starts from the next page.

Älä muuta mitään lausetta sen jälkeen, kun olet kääntänyt sen. Minua kiinnostaa ensimmäinen yrityksesi kääntää nämä lauseet. Sinulla on aikaa 15 minuttia. Yritä kääntää siinä ajassa kaikki lauseet.

Jos sinulla ei ole kysymyksiä, voit aloittaa. Onnea matkaan!

* * * * *

Käännöstesti alkaa seuraavalla sivulla

1. Lasi hajoaa, kun putoaa lattialle.

glass	break	whenever	fall	floor
-------	-------	----------	------	-------

.....

2. Tunsin itseni väsyneeksi kun pääsin kotiin.

feel	tired	when	get	home
------	-------	------	-----	------

.....

3. Jane sanoo, että rakastaa sinua.

say	love
-----	------

.....

4. Äitini istuu tuolissaan aina, kun on väsynyt.

mother	sit	Chair	whenever	feel	tired
--------	-----	-------	----------	------	-------

.....

5. Hän hukkasi kirjansa, kun jätti sen junaan

lose	book	When	leave	train
------	------	------	-------	-------

.....

6. Hän sanoo, että haluaa ostaa uuden auton

say	want	buy	car
-----	------	-----	-----

.....

7. Hän soittaa minulle, kun saapuu.

call	When	arrive
------	------	--------

.....

8. He eivät lähde toimistosta, ennen kuin tapaavat johtajan.

leave	office	Until	meet	manager
-------	--------	-------	------	---------

.....

9. John väittää, että puhuu englantia hyvin.

claim	Speak	English
-------	-------	---------

.....

10. John mursi jalkansa, kun hyppäsi puusta

break	leg	when	jump	tree
-------	-----	------	------	------

.....

11. Pojat olivat onnellisia, kun voittivat palkinnon.

boy	happy	when	win	prize
-----	-------	------	-----	-------

.....

12. Mary piti kissasta, kun näki sen.

like	cat	when	see
------	-----	------	-----

.....

13. Sanoin, että löysin kynäni.

say	Find	pen
-----	------	-----

.....

14. Luulet, että jätit kirjasi luokkaan

think	leave	book	class
-------	-------	------	-------

.....

15. Et voi katsoa elokuvaa TV:stä kunnes syöt ruokasi loppuun.

watch	movie	Until	finish	food
-------	-------	-------	--------	------

.....

3. Grammatical and Ungrammatical English sentences

Instructions

Here are some sentences. Some are written correctly in English and others contain an error. Please read the sentences carefully and then tick one of the options:

- CLEARLY CORRECT English sentence
- CLEARLY INCORRECT English sentence
- POSSIBLY INCORRECT English sentence
- I DON'T KNOW

Here is an example:

(i) Mr Smith's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Please tick: CLEARLY CORRECT if you are sure that the sentence is GRAMMATICAL (written correctly/ has no error) in English.

2. Kieliopillisesti oikeat ja väärät englanninkieliset lauseet

Ohjeet

Tässä on muutamia lauseita. Jotkut ovat oikein englanniksi, ja joissain on virhe. Lue lauseet huolellisesti ja valitse yksi vaihtoehdoista:

- CLEARLY CORRECT (**SELVÄSTI OIKEIN englanniksi**)
- CLEARLY INCORRECT (**SELVÄSTI VÄÄRIN englanniksi**)
- POSSIBLY INCORRECT (**MAHDOLLISESTI VÄÄRIN englanniksi**)
- I DON'T KNOW (**EN TIEDÄ**)

Esimerkiksi:

(i) Mr Smith's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Valitse: CLEARLY CORRECT (SELVÄSTI OIKEIN), jos olet varma siitä, että lause on KIELIOPILLISESTI OIKEIN (virheetön/kirjoitettu oikein) englanniksi.

Please tick: CLEARLY INCORRECT if you are sure that the sentence is UNGRAMMATICAL (has an error).

Please tick: POSSIBLY INCORRECT if you think there is an error but you are **not certain** about it.

Please tick: I DON'T KNOW if you have **no** idea about the answer.

If you tick **INCORRECT** or **POSSIBLY INCORRECT**, please **draw a line** under the part of the sentence that you believe is wrong. Then, if possible, please correct the mistake by rewriting only the wrong part of the sentence. Look at the example below: If you think that the sentence is Incorrect because of the missing article in 'gave him gift', you would draw a line and add an 'a' as follows:

(ii) Mr Smith's students gave him gift at the end of the year.

Clearly correct

Clearly incorrect

I don't know

Possibly incorrect

It should be: gave him a gift

Valitse: CLEARLY INCORRECT (SELVÄSTI VÄÄRIN), jos olet varma siitä, että lause on KIELIOPILLISESTI VÄÄRIN (siinä on virhe).

Valitse: POSSIBLY INCORRECT (MAHDOLLISESTI VÄÄRIN), jos uskot että lauseessa saattaa olla virhe, mutta et ole varma.

Valitse: I DON'T KNOW (EN TIEDÄ), jos et tiedä vastausta.

Jos valitset vaihtoehdon **INCORRECT** (VÄÄRIN) tai **POSSIBLY INCORRECT** (MAHDOLLISESTI VÄÄRIN), **alleviivaa** se osa lauseesta, jonka uskot olevan väärin. Jos mahdollista, korjaa virhe kirjoittamalla uudelleen vain virheellinen osa lauseesta. Katso alla olevaa esimerkkiä: jos uskot lauseen olevan VÄÄRIN, sillä kohdasta 'gave him gift' puuttuu artikkeli, alleviivaisit ja lisääisit kirjaimen 'a' seuraavasti:

(ii) Mr Smith's students gave him gift at the end of the year.

Clearly correct

Clearly incorrect

I don't know

Possibly incorrect

Sen pitäisi olla: gave him a gift

Important Notes

1. Remember: If you think the sentence could be **incorrect**, choose POSSIBLY INCORRECT and change the sentence as in the example of (ii).
2. You have 40 minutes for this test; please make sure you finish all the items.
3. Remember, this test is for research purposes, so you do not need to worry about your score.

Here are three examples to try before you start the test:

(iii) I need to borrow my sister's Finnish-English dictionary.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

(vi) Bill wondered where is Mary going shopping.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: _____

Tärkeitä huomautuksia

1. Huomaa: Jos mielestäsi lauseessa saattaa olla virhe, älä valitse vaihtoehtoa I DON'T KNOW (EN TIEDÄ). Valitse sen sijaan POSSIBLY INCORRECT (MAHDOLLISESTI VÄÄRIN) ja muuta lausetta kuten esimerkissä (ii).
2. Sinulla on aikaa 40 minuuttia. Tarkoitus on että suoritat siinä ajassa kaikki tehtävät.
3. Muista: Tämä testi on vain tutkimusta varten. Älä siis huolehdi tuloksestasi

Tässä on kolme esimerkkiä, joita voit kokeilla ennen testin aloittamista:

(iii) I need to borrow my sister's Finnish-English dictionary.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Sen pitäisi olla: _____

(vi) Bill wondered where is Mary going shopping.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

Sen pitäisi olla: _____

(v) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

(See the answers below)

Answers

(iii) I need to borrow my sister's Finnish-English dictionary.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

(iii) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: hopes the dictionary

(v) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect.

Sen pitäisi olla: _____

(Katso alla olevat vastaukset)

Vastaukset

(iii) I need to borrow my sister's Finnish-English dictionary.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

Sen pitäisi olla: _____

(iii) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

Sen pitäisi olla: hopes the dictionary

(vi) Bill wondered where is Mary going shopping.

Clearly correct

Clearly incorrect

I don't know

Possibly incorrect

It should be: Mary is.

Do you have questions before you start this test? If you have, please ask the researcher right now. Otherwise, you are now ready to begin. Have fun!

The test starts from the next page.

NB: See Appendix 1.a above for a complete listing of the GJ test sentences

(vi) Bill wondered where is Mary going shopping.

Clearly correct

Clearly incorrect

I don't know

Possibly incorrect

Sen pitäisi olla: Mary is.

Onko sinulla kysyttävää ennen tämän testin aloittamista? Jos sinulla on kysyttävää, käänny nyt tutkijan puoleen. Jos sinulla ei ole kysyttävää, voit aloittaa. Onnea matkaan!

Testi alkaa seuraavalla sivulla

Appendix 1.c. The GJ task and the translation task
(Arabic version)

Study investigating knowledge of English

4. The Translation Task

Instructions

Here are some sentences in your own language. Please translate **all** of them into English. After each sentence there are some suggestions of words that you could use in your translation. If you do not understand a word, please look at the vocabulary list (it has words with their meanings in Arabic). **Please follow these models:**

1. قابلت فاطمة هند أمس.

Meet	yesterday
------	-----------

Model Answer: Fatima met Hind yesterday

2. يحب الأولاد اللعب في الحديقة.

Boy	like	play	park
-----	------	------	------

Model Answer: The boys like playing in the park.

دراسة لتقويم مدى إلمام الطالب باللغة الإنجليزية

1- اختبار الترجمة

التعليمات

تجد أدناه بعض الجمل بلغتك. الرجاء ترجمتها الى اللغة الإنجليزية. و بعد كل جملة ستجد بعض الكلمات التي يمكن أن تساعدك في ترجمة الجملة إلى اللغة الإنجليزية, وفي حال عدم إلمامك بأي من هذه المفردات يمكنك الرجوع إلى قائمة المفردات/الكلمات المرفقة مع معانيها باللغة العربية.

إليك المثالين التاليين:

1. قابلت فاطمة هند أمس.

Meet	yesterday
------	-----------

(إجابة نموذجية) : Fatima met Hind yesterday

2. يحب الأولاد اللعب في الحديقة.

boy	like	play	park
-----	------	------	------

(إجابة نموذجية) : The boys like playing in the park

Please do **not** change any sentence after you have translated it. I am interested in your first attempt at translating these sentences. You will have 15 minutes on this exercise; please try to complete all of the sentences.

If you have no questions, please start. Have fun!

* * * * *

The translation test starts from the next page.

الرجاء عدم تغيير أي جملة بعد ترجمتها. أنا مهتم بمحاولتك الأولى لترجمة الجمل. الرجاء ترجمة جميع الجمل التالية في وقت لا يتجاوز 15 دقيقة.

إذا لم يكن لديك أي أسئلة فأرجو البدء في الاختبار، أتمنى لك وقتاً

ممتعاً!

* * * * *

الاختبار يبدأ من الصفحة التالية (صفحة رقم 3)

1. ينكسر الزجاج عندما يسقط على الأرض.

glass	break	When	fall	floor
-------	-------	------	------	-------

2. شعرت بتعب شديد عندما وصلت إلى المنزل.

feel	tired	when	get	home
------	-------	------	-----	------

3. تقول امي إنها تحبك.

say	love
-----	------

4. تجلس والدتي على الكرسي كلما شعرت بالتعب.

mother	sit	chair	whenever	feel	tired
--------	-----	-------	----------	------	-------

5. فقدت كتابها عندما تركته في القطار.

lose	book	When	leave	train
------	------	------	-------	-------

6. يقول إنه يريد شراء سيارة جديدة.

say	want	buy	car
-----	------	-----	-----

7. قال إنه سوف يتصل بي عندما يصل.

call	When	arrive
------	------	--------

8. لن يغادروا المكتب حتى يقابلوا المدير.

leave	office	Until	meet	manager
-------	--------	-------	------	---------

9. يدعي محمد أنه يتحدث الإنجليزية جيدا .

claim	Speak	English
-------	-------	---------

10. كسر أحمد ساقه عندما قفز من الشجرة .

break	leg	when	jump	tree
-------	-----	------	------	------

11. كان الأولاد سعداء عندما فازوا بالجائزة .

boy	happy	when	win	prize
-----	-------	------	-----	-------

12. أحببت هند القطة عندما رأته .

like	cat	when	see
------	-----	------	-----

13. قلت إنني وجدت قلمي .

say	Find	pen
-----	------	-----

14. تعتقد أنك نسيت كتابك في المدرسة .

think	forget	book	school
-------	--------	------	--------

15. لا يمكنك مشاهدة التلفاز حتى تنهي طعامك .

watch	movie	Until	finish	food
-------	-------	-------	--------	------

2- الجمل الإنجليزية الصحيحة لغويا وغير الصحيحة

التعليمات

إليك بعض الجمل بلغة إنجليزية، بعض هذه الجمل صحيحة وبعضها تحتوي على أخطاء. الرجاء قراءة كل جملة ثم اختر واحدا من هذه الخيارات:

- جملة إنجليزية صحيحة (CLEARLY CORRECT)
- جملة إنجليزية غير صحيحة (CLEARLY INCORRECT)
- جملة إنجليزية من الممكن أن تكون غير صحيحة (POSSIBLY INCORRECT)
- لا أعلم (I DON'T KNOW)

إليك مثال :

4. Grammatical and Ungrammatical English sentences

Instructions

Here are some sentences. Some are written correctly in English and others contain an error. Please read the sentences carefully and then tick one of the options:

- **CLEARLY CORRECT** English sentence
- **CLEARLY INCORRECT** English sentence
- **POSSIBLY INCORRECT** English sentence
- **I DON'T KNOW**

Here is an example:

(i) Mr Smith's students gave him gift at the end of the year.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

Please tick: *CLEARLY CORRECT* if you are sure that the sentence is GRAMMATICAL (written correctly/ has no error) in English.

(i) Mr Smith's students gave him gift at the end of the year.

- Clearly correct Clearly incorrect
 Possibly incorrect I don't know.

ضع إشارة (✓) بجانب (CLEARLY CORRECT) إذا كنت متأكدًا أن الجملة صحيحة و مكتوبة بقواعد سليمة باللغة الإنجليزية.

Please tick: CLEARLY INCORRECT if you are sure that the sentence is **UNGRAMMATICAL** (has an error).

Please tick: POSSIBLY INCORRECT if you think there is an error but you are **not certain** about it.

Please tick: I DON'T KNOW if you have **no** idea about the answer.

If you tick **INCORRECT** or **POSSIBLY INCORRECT**, please **draw a line** under the part of the sentence that you believe is wrong. Then, if you can, please correct the mistake by rewriting only the wrong part of the sentence. Look at the example below: If you think that the sentence is *incorrect* because of the missing article in 'gave him gift', you would draw a line and add an 'a' as follows:

(ii) Mr Smith's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input checked="" type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

It should be: gave him a gift

ضع إشارة (✓) بجانب (CLEARLY INCORRECT) إذا كنت متأكداً أن الجملة غير صحيحة من الناحية اللغوية.

ضع إشارة (✓) بجانب (POSSIBLY INCORRECT) إذا كنت تعتقد أن الجملة قد يكون فيها خطأ لغوي ولكنك غير متأكد منه.

ضع إشارة (✓) بجانب (I DON'T KNOW) إذا لم يكن لديك أي فكرة عن الإجابة.

إذا وضعت إشارة (✓) بجانب (CLEARLY INCORRECT) أو بجانب (POSSIBLY INCORRECT) الرجاء وضع خط تحت الجزء الذي فيه خطأ في الجملة، ثم الرجاء تصحيح الخطأ وإعادة كتابة الجزء الخاطئ فقط من الجملة.

انظر إلى المثال أدناه. إذا كنت تعتقد أن الجملة غير صحيحة بسبب عدم وجود أداة تعريف في عبارة "gave him gift" فسوف تضع خط تحت العبارة وتضيف "a" كما يلي:

(ii) Mr Smith's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly correct | <input checked="" type="checkbox"/> Clearly incorrect |
| <input type="checkbox"/> I don't know | <input type="checkbox"/> Possibly incorrect |

يفترض أن تكون: gave him a gift

Important Notes

1. Remember: If you think the sentence **could be incorrect**, choose **POSSIBLY INCORRECT** and change the sentence as in example (ii).
2. You have 40 minutes for this test; please make sure you finish all the items.
3. Remember, this test is for research purposes, so you do not need to worry about your score.

Here are three examples to try before you start the test:

(iii) I need to borrow my sister's Arabic-English dictionary.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

(iv) Bill wonders where is Mary going shopping.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

ملاحظات مهمة :

- 1- تذكر: إذا كنت تعتقد أن الجملة يمكن أن تكون خاطئة يجب عليك ان تختار **(POSSIBLY INCORRECT)** ، ومن ثم صحح الجمل كما هو موضح في المثال السابق (ii).
- 2- الرجاء الاجابة عن جميع الجمل في وقت لا يتجاوز 40 دقيقة.
- 3- لا تقلق حيال درجتك في هذا الاختبار لأنه صمم لغرض بحثي. إليك ثلاثة أمثلة للتجربة قبل البدء في الامتحان:

(iii) I need to borrow my sister's Arabic-English dictionary.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

يفترض أن تكون: _____

(iv) Bill wonders where is Mary going shopping.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

يفترض أن تكون: _____

(v) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

(See the answers below)

Answers

(iii) I need to borrow my sister's Arabic-English dictionary.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: _____

(vi) Bill wonders where is Mary going shopping.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: Mary is.

(v) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

_____ يفترض أن تكون:

انظر الإجابات فيما يلي:

الإجابات

(iii) I need to borrow my sister's Arabic-English dictionary.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

_____ يفترض أن تكون:

(vi) Bill wonders where is Mary going shopping.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

Mary is. يفترض أن تكون:

(v) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

It should be: hopes the dictionary

Do you have questions before you start this test? If you have, please ask the researcher right now. Otherwise, you are now ready to begin. Have fun!

The test starts from the next page.

NB: See Appendix 1.a above for a complete listing of the GJ test sentences

(v) Mary hopes dictionary she bought will help her to improve her English.

- Clearly correct Clearly incorrect
 I don't know Possibly incorrect

hopes the dictionary: يفترض أن تكون

هل لديك أسئلة قبل البدء الاختبار؟ يمكنك أن تسأل الباحث الآن اذا كان لديك أي سوال.

وإلا فتفضل وابدأ الاختبار الآن، استمتع بوقتك.

يبدأ الاختبار من بداية الصفحة التالية (صفحة رقم 10)

Appendix 2: Lists of English words with their meanings in the source languages

Appendix 2.a. French version

abroad	à l'étranger
allow	permettre
arrive	arriver
before	avant que
believe	croire
chase	chasser
claim	prétendre
complain	se plaindre
cook	cuisiner
deliver	livrer
escape	s'échapper
follow	suivre
frequently	souvent
hit	frapper
jump	sauter
mark	note
marry	se marier
mention	dire
nearly	presque
noise	bruit
prize	prix
race	course
return	revenir
situation	situation
steal	voler
town	ville
unhappy	malheureux, malheureuse
until	jusqu'à ce que
vase	vase
when	quand
whenever	quand (chaque fois qu')

Appendix 2.b. Finnish version

abroad	uikomailla
allow	sallia
arrive	saapua
before	ennen kuin
believe	uskoa
chase	ajaa takaa
claim	vaatia
complain	valittaa
cook	keittää
deliver	jakaa
escape	paeta
follow	seurata
frequently	usein
hit	Lyödä
insist	väittää
jump	hypätä
mark	arvosana
marry	mennä naimisiin
mention	mainita
nearly	lähes
noise	melu
prize	palkinto
race	kilpailu
return	palata
situation	tilanne
steal	varastaa
town	kaupunki
unhappy	onneton
until	ennen kuin / kunnes
vase	vaasi
when	kun
whenever	aina kun

Appendix 2.c. Arabic version

abroad	خارج البلاد
allow	يسمح
arrive	يصل
before	قبل
believe	يعتقد
chase	يطارد
claim	يدعي
complain	يشكو / يتذمر
cook	يطبخ
deliver	يوصل
escape	يهرب
follow	يتبع / يقتفي
frequently	بشكل متكرر
hit	يصدم
jump	يقفز
mark	درجة / علامة مدرسية
marry	يتزوج
mention	يذكر / يشير الى
nearly	على وشك / تقريبا
noise	ضجيج
prize	جائزة
race	سباق
return	يعود
situation	وضع / حالة
steal	يسرق
town	بلدة / مدينة
unhappy	غير سعيد
until	حتى
vase	مزهريّة
when	حين / عندما
whenever	كلما

Appendix 3. Initial pilot experimental test version

Study investigating knowledge of English

1. The Translation Task

Instructions

In the following task you will find a series of sentences written in your own language. Please translate **all** of them into English. Each sentence is followed by some words that you may use to improve your ability to translate the sentences.

Please follow these models:

1. نسيت قلمي.

forget	pen
--------	-----

I forgot my pen.

2. ذهب احمد الى فرنسا الصيف الماضى.

Go	France	last	Summer
----	--------	------	--------

Ahmed went to France last summer.

Please do **not** go back to edit any sentence after translating it. I am interested in your first attempt at translating these sentences. Hence, it is **essential** that you do **not** take a lot of time in doing this task, as spending too much time in this task will invalidate my results. So, you **should not** spend more than 20 minutes to complete this task.

If you don't have any question, please start. Have fun!

The test starts here

1. يحب أبي قراءة الصحف في الصباح قبل ذهابه للعمل.

Like	father	read	newspapers	morning	before	go	Work
------	--------	------	------------	---------	--------	----	------

2. كسرت ساق أحمد عندما قفز من النافذة.

break	Ahmed	leg	When	jump	window
-------	-------	-----	------	------	--------

3. أستطيع مقابلتك اليوم بعد أن أنتهي من عملي.

can	Meet	today	after	finish	work
-----	------	-------	-------	--------	------

4. شعرت بتعب شديد عندما عدت للمنزل ليلة البارحة.

feel	Tired	when	Get	home	Last	Night
------	-------	------	-----	------	------	-------

5. قالت فاطمة إنها تريد شراء منزل جديد.

say	Fatima	want	Buy	new	house
-----	--------	------	-----	-----	-------

6. يبكي طفلي كلما سمع صوت الرعد.

cry	child	whenever	Hear	thunder
-----	-------	----------	------	---------

7. قال أحمد إن الاطفال يعتقدون أنه ذاهب إلى المستشفى.

say	Ahmed	child	think	go	hospital
-----	-------	-------	-------	----	----------

8. يزور أخي والدتنا في المستشفى كل يوم بعد إنهاء عمله.

visit	brother	mother	hospital	Everyday	after	finish	Work
-------	---------	--------	----------	----------	-------	--------	------

9. توقفت السيارة قبل أن تصطدم بالولد.

stop	car	Before	Hit	Boy
------	-----	--------	-----	-----

10. أخبر المعلم طلابه أنهم جميعا قد نجحوا.

inform	teacher	student	Pass	exam
--------	---------	---------	------	------

2. Grammatical and Ungrammatical English sentences

Instructions and Procedures

In this task you will see a set of sentences. Some are perfectly grammatical while others contain an error. Please read the sentences carefully, and then following each item tick one of the options: CLEARLY GRAMMATICAL English sentence, POSSIBLY UNGRAMMATICAL English sentence, CLEARLY UNGRAMMATICAL English sentence, or I DON'T KNOW. Follow the following examples:

(i) Mr John's students gave him a gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly grammatical | <input type="checkbox"/> Possibly ungrammatical |
| <input type="checkbox"/> Clearly ungrammatical | <input type="checkbox"/> I don't know. |

You will tick: Grammatical English Sentence , if you are sure that the sentence is GRAMMATICAL in English and you could say it under appropriate circumstances (e.g., whether in casual or in formal conversation). However, if you read the sentence:

(ii) Mr John's students gave him gift at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly grammatical | <input type="checkbox"/> Possibly ungrammatical |
| <input type="checkbox"/> Clearly ungrammatical | <input type="checkbox"/> I don't know. |

If you are confident that this sentence is ungrammatical and you would **never** say it under any circumstance, please tick: Ungrammatical English Sentence . However, if you doubt you would say it under any circumstances, but you think it could be grammatical, please tick: POSSIBLY UNGRAMMATICAL .

If you believe that the sentence is ungrammatical or is possibly ungrammatical, please **draw a pen or pencil line** under the part of the sentence that you believe makes it ungrammatical/possibly ungrammatical. For example, in the case of (ii), you would draw a line under 'gave him gift' as follows:

(ii) Mr John's students **gave him gift** at the end of the year.

- | | |
|--|---|
| <input type="checkbox"/> Clearly grammatical | <input type="checkbox"/> Possibly ungrammatical |
|--|---|

Clearly ungrammatical

I don't know.

The reason is that this sentence gave him gift is not acceptable in English. It should be: Mr John's students gave him a gift at the end of the year.

If you **totally** do not know, **do not guess**, but choose DON'T KNOW option. Remember **do not** tick the DON'T KNOW option if you have doubts about the sentence but **choose** instead the POSSIBLY UNGRAMMATICAL option.

Also, please do not judge these based on preference. That is, if the sentence is GRAMMATICALLY POSSIBLE please **do not** reject it simply because you would prefer to say it in a different manner. When you judge sentences, you must focus on whether or not a sentence is a possible English sentence, not on whether you prefer to say it in a different way because there is always more than one way to express the same meaning.

Last but not least, it is **important** that you **do not** spend a lot of time trying to find out what language rules might be violated in the sentence based. I am interested in your **immediate** reaction to the sentences – your initial feeling towards the sentences. Therefore, this test **should be completed** in no more than 35 minutes.

Do you have questions before you start the task? If you have, please ask the researcher right now. Otherwise, you are now ready to begin. Have fun!

The test starts from the next page

1. I will lend you my grammar book, if you will give it back to me before class tomorrow.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

2. Sara claims that this morning Jane hit with a book.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

3. She will never marry until finds the right man.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

4. John told me that his mother often drink tea after dinner.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

5. It was the hottest summer that Jane had ever known.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

6. My father likes to read the newspaper in the morning before gets to the office.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

7. John may come if comes alone
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

8. Jim lost his iPhone and asked the police to find for him
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

9. Michael thought that the cinema often shows films in the afternoons during the winter.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
10. My father visits his mother in hospital every day when finishes his work.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
11. Mary suggested that the park always close too early in the summer.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
12. The boys were unhappy when came last in the race.'
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
13. Bill complains that yesterday the teacher stopped from playing football during the lunch hour.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
14. John could get fired if he miss any morning meetings.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
15. Thomas broke the same leg he had broken last summer when jumped from the tree.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.

16. I can see you when finish my work.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
17. Sam said that the postman frequently deliver letters to the wrong house.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
18. I felt very tired when got home late last night.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
19. My brother bought the golden ring because his wife loved.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
20. The little girls spent several hours playing in the park.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
21. Last night I had just gone to bed when heard a noise in the kitchen.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
22. I am pleased that my child like his English teacher.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
23. In the park the gardener was working when slipped and hurt his head.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

24. Sara told me that her father often reads the newspaper after breakfast.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
25. John and Ben report that yesterday the teacher allowed to miss the math class to play football.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
26. Mary said that John thinks that he is smart.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
27. I had nearly finished my book when left it on the train.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
28. I was putting my shoes on when fell over suddenly.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
29. You should lock your bicycle to something in case somebody tries to steal it.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
30. Jane believed that the University sometimes give scholarships to very good students.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
31. My child often cries when hears thunder.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.

32. Bill said that the library usually open on Sundays during the summer.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
33. The car stopped before hit the child.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
34. Sally comes back to say sorry for the cup she broke earlier.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
35. When I have to catch the train, I am always worried that I will miss.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
36. Bill says that the children think that went to see a doctor.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
37. Mary hopes it will not be too long before the bus arrive so that she can finally get home.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
38. Mary likes playing in the park with her friends where she picks flowers and takes them home to give them to her mother.
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.
39. She had just fixed her bicycle when broke again
- Clearly grammatical Possibly ungrammatical
- Clearly ungrammatical I don't know.

40. The professor informed Mary that passed the exam.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
41. Jane may come if she come alone
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
42. She enjoyed the trip but her eyes hurt so much that she started to cry.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
43. If you do not want that sandwich, threw to the birds to eat it.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
44. Jane hopes it will not be too long before the train arrives.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
45. I was driving my car in a highway when realised that I was nearly out of fuel.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
46. He said nothing in reply although he was deeply hurt by her remarks.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.
47. Mary says that wants to buy a new house.
 Clearly grammatical Possibly ungrammatical
 Clearly ungrammatical I don't know.

48. I have not seen him since he returned to the country.
- | | |
|--|---|
| <input type="checkbox"/> Clearly grammatical | <input type="checkbox"/> Possibly ungrammatical |
| <input type="checkbox"/> Clearly ungrammatical | <input type="checkbox"/> I don't know. |
49. I phoned Mary and invited to the party
- | | |
|--|---|
| <input type="checkbox"/> Clearly grammatical | <input type="checkbox"/> Possibly ungrammatical |
| <input type="checkbox"/> Clearly ungrammatical | <input type="checkbox"/> I don't know. |
50. Ann was in a very difficult situation when lost her job.
- | | |
|--|---|
| <input type="checkbox"/> Clearly grammatical | <input type="checkbox"/> Possibly ungrammatical |
| <input type="checkbox"/> Clearly ungrammatical | <input type="checkbox"/> I don't know. |

Thank you

Appendix 4. Consent forms

Appendix 4.a. Version for the adult participants



CONSENT FORM

TO BE COMPLETED BY ADULT PARTICIPANTS

A. Purpose and Background of this Research Study

You are invited to participate in a research study that investigates the intuitive knowledge of English speakers — both native and non-native speakers — to discover what sort of language is natural for most people.

B. Procedures

If you agree to participate, you are asked to perform a short translation task of approximately 15 minutes and a grammaticality (grammar) judgment task (GJT) of approximately 40 minutes. In the first task, you will translate sentences from your native language into English. In the second task, you will rate some sentences. If you are a native speaker of English, you will only do the second task.

C. Risks

There are no foreseeable risks associated with this study. However, if you feel discomfort at any time, notify the researcher and you can discontinue the study.

D. Direct Benefits

You will receive £8 or the equivalent amount in euros or Saudi Riyals as payment for your participation.

E. Anonymity and Confidentiality

All information in this study will be anonymous. Your identity will be kept confidential; your name will not be used in reports on the data.

All data will be stored securely either electronically or in a locked cupboard, and no one other than the researcher will have access to the data. As part of the data analysis process, hard copies of surveys (without names attached) may be given to the doctoral supervision team or a small number of other researchers. Hard copies will be kept by the researcher.

F. Consent

I agree to take part in the above project/investigation. I have read and understood the information provided above. I know what the study is about. I know that my participation is voluntary and I have the right to withdraw from the study at any point.

Name _____

Signature _____ **Date** ____ / ____ / **2013**

Investigator:

Naif Alsaedi, post-graduate student at Newcastle University, UK

Email: n.s.alsaedi@newcastle.ac.uk

Supervisors:

Prof Martha Young-Scholten, Professor of Applied Linguistics, Newcastle University

Email: martha.young-scholten@ncl.ac.uk

Prof Anders Holmberg, Professor in Theoretical Linguistics, Newcastle University

Email: anders.holmberg@ncl.ac.uk

Appendix 4.b. Version for the minor participants



CONSENT FORM

TO BE COMPLETED BY STUDY PARTICIPANT AND PARENT/GUARDIAN

A. Purpose and Background of this Research Study

You are invited to participate in a research study that investigates the intuitive knowledge of English speakers — both native and non-native speakers — to discover what sort of language is natural for most people.

B. Procedures

If you agree to participate, you will be asked to perform a short translation task of approximately 15 minutes and a grammaticality (grammar) judgment task (GJT) of approximately 40 minutes. In the first task, you will translate sentences from your native language into English. In the second task, you will be asked to rate some sentences. If you are a native speaker of English, you will only do the second task.

C. Risks

There are no foreseeable risks associated with this study. However, if you feel discomfort at any time, notify the researcher and you can discontinue the study.

D. Direct Benefits

You will receive £8 or the equivalent amount in euros or Saudi Riyals as payment for your participation.

E. Anonymity and Confidentiality

All information in this study will be anonymous. Your identity will be kept confidential; your name will not be used in reports on the data.

All data will be stored securely either electronically or in a locked cupboard; no one other than the researcher will have access to the data. As part of the data analysis process, hard copies of surveys (without names attached) may be reviewed by the doctoral supervision team or a small number of other researchers. Hard copies will be kept by the researcher.

F. Consent

PART A: TO BE COMPLETED BY THE PARTICIPANT

I agree to take part in the above project/investigation. I have read and understood this form. I know what the study is about. I know that my participation is voluntary and I have the right to withdraw from the study at any point.

Name _____

Signature _____ **Age** _____

PART B: TO BE COMPLETED BY THE PARENT/GUARDIAN

I have read and understood the accompanying letter and give permission for the child (named above) to be included.

Name _____

Relationship to child _____

Investigator:

Naif Alsaedi, post-graduate student at Newcastle University, UK
Email: n.s.alsaedi@newcastle.ac.uk

Supervisors:

Prof Martha Young-Scholten, Professor of Applied Linguistics, Newcastle University
Email: martha.young-scholten@ncl.ac.uk

Prof Anders Holmberg, Professor in Theoretical Linguistics, Newcastle University
Email: anders.holmberg@ncl.ac.uk

Appendix 5. Personal information form

Study investigating knowledge of English

Thank you for agreeing to participate in this study, which investigates the intuitive knowledge of speakers of English to discover what sort of language is natural for most people.

To enable me to consider the relevance of factors like age, sex, knowledge of other languages etc., I have a few questions for you. As already mentioned in the consent form, all information in this study will be anonymous. Your identity will be kept confidential; your name will not be used in reports on the data.

1. Your name (optional):
2. Your email (optional):
3. Your native language(s):
4. Are you: female male
5. Your age or date of birth:
6. Age at which you first started learning English (write 'native' if you are a native speaker, and go to question 9):
7. Number of years you have attended English classes:
8. Number of months you have lived in an English-speaking community:
.....
9. Other languages you speak fluently:
10. Other languages you speak moderately:

Please note that participating in this study **requires** unfamiliarity with linguistics as an academic subject.

The investigator is:

Naif Alsaedi, post-graduate student at Newcastle University – UK.

Email: n.s.alsaedi@newcastle.ac.uk

DEBRIEFING FORM

Study title: **The Use of Pronominal Subjects by Arabic, Finnish and French Speakers of English**

Thank you for your participation in this study. The answers that you provided on the grammar ask and translation task will help the researcher answer specific questions related to second language acquisition.

The general purpose of this research is to investigate the acquisition of the obligatory pronominal subjects (the use of pronouns) in English by native speakers of Arabic, Finnish and French. It tests whether those learners transfer the grammar of their first language to English or whether they follow a developmental path similar to that of all second language learners when acquiring a foreign language. Your responses help address the issue of linguistic transfer, meaning the parts of speech that carry over to a new language (English in this case) or that are generalized from a person's native language into the new language.

If you have any questions about this study, contact the researcher Mr Naif Alsaedi, School of English Literature, Language and Linguistics, Percy Building, Newcastle University, Newcastle upon Tyne NE1 7RU, England. Telephone: +44 (0) 758 652 3054. Email: n.s.alsaedi@newcastle.ac.uk

If you have any additional concerns about any aspect of the study, you may contact:

Prof Martha Young-Scholten, Professor of Applied Linguistics, Newcastle University [martha.young-scholten@ncl.ac.uk]

Prof Anders Holmberg, Professor in Theoretical Linguistics, Newcastle University [anders.holmberg@ncl.ac.uk]

Thank you again for your assistance.

Naif Alsaedi

Appendix 7. Oxford Online Placement Test



Language Centre

Name:

.....

Date:

.....

Quick Placement Test

Online version

http://www.lang.ox.ac.uk/courses/tst_placement_english.html

Answer all of the following questions below by selecting an answer from the list.

1. How many people _____ in your family?

- A) are they
- B) is it
- C) are there
- D) is

2. What time is it? _____

- A) Ten and a quarter.
- B) Ten minus quarter.
- C) A quarter past ten.
- D) Fifteen after ten o'clock.

3. I get up at 8 o'clock _____ morning.

- A) in the
- B) in
- C) the
- D) at the

4. How much _____ where you live?

- A) do houses cost
- B) does houses cost
- C) does cost houses
- D) do cost houses

5. Where are you going _____ Friday?

- A) at
- B) in
- C) on
- D) the

6. _____ come to my party next Saturday?

- A) Do you can
- B) Can you to
- C) Can you
- D) Do you

7. What _____ in London last weekend?

- A) you were doing
- B) did you do
- C) you did
- D) did you

8. Is your English improving? _____

- A) I hope it.
- B) Hoping.
- C) I hope so.
- D) I hope.

9. I'm going to Sainsbury's _____ some food.

- A) buy
- B) for buy
- C) to buy
- D) for to buy

10. Oxford is the most attractive city _____

- A) I've ever seen.
- B) that I see.
- C) I've never seen.
- D) that I saw already.

11. Oxford isn't _____ Bath.

- A) as beautiful than
- B) so beautiful than
- C) so beautiful that
- D) as beautiful as

12. He was mowing the lawn when I _____ him yesterday.

- A) saw
- B) had seen
- C) was seeing
- D) have seen

13. Last Tuesday I _____ to the Passport Office.

- A) must gone
- B) must go
- C) had to go
- D) had go

14. What were you doing at 7:30 on Wednesday evening? I _____ TV.

- A) was watching
- B) watched
- C) was watched
- D) watching

23. My mother never _____ out in the evenings.

- A) goes
B) go
C) is going
D) going

24. _____ Oxford?

- A) Since when you live in
B) How much time you are living in
C) How long have you been living in
D) How long time are you living in

25. _____ car is the red Ford?

- A) Whose
B) To whom
C) Who's
D) Of who

26. I'm sorry. I haven't done my report _____

- A) up to the now.
B) already.
C) until the present.
D) yet.

27. My friend doesn't speak Chinese. I don't _____

- A) also.
B) neither.
C) either.
D) too.

28. That's the house _____

- A) in the which Mr Brown lives.
B) in which Mr Brown lives in that.
C) Mr Brown lives in.
D) Mr Brown lives in that.

29. If _____

- A) you come to my office, I'd pay you.
B) you shall come to my office, I'll pay you.
C) you come to my office, I would to pay you.
D) you come to my office, I'll pay you.

30. She asked me how big _____
 A) is your house. B) my house was.
 C) was my house. D) is my house.
31. My friend let _____ his bike yesterday.
 A) to borrow B) me borrowing
 C) me to borrow D) me borrow
32. _____, what would you spend it on?
 A) When you had a lot of money B) If you had a lot of money
 C) If you would have a lot of money D) If you shall have a lot of money
33. I _____ smoking last year, but I didn't.
 A) ought to give up B) ought to have given up
 C) ought given up D) ought to give up
34. I'm _____ the film on Wednesday.
 A) looking forward to see B) looking forward to seeing
 C) look forward seeing D) looking forward seeing
35. I'm not _____ grammar.
 A) interested to learn B) interested in learning
 C) interesting to learning D) interesting in learning
36. The film was very good. It's _____
 A) worth seeing. B) worth to see.
 C) worthwhile to see. D) worthwhile see.
37. I have difficulty _____ English.
 A) to write B) writing
 C) about writing D) to writing

Appendix 8. List of Shapiro Results

The GJ Task

Table 1. Null-subjects acceptance

L1	level	p-value
French	LI	0.494
	UI	0.056
	ADV	0.000
Finnish	LI	0.204
	UI	0.030
	ADV	0.000
Arabic	LI	0.390
	UI	0.668
	ADV	0.030

Table 2. Null-subjects acceptance
(in adverbial clauses)

L1	level	p-value
French	LI	0.090
	UI	0.107
	ADV	0.000
Finnish	LI	0.104
	UI	0.034
	ADV	0.000
Arabic	LI	0.263
	UI	0.826
	ADV	0.025

Table 3. Null-subjects acceptance
(in complement clauses)

L1	level	p-value
French	LI	0.003
	UI	0.000
	ADV	0.000
Finnish	LI	0.425
	UI	0.000
	ADV	0.000
Arabic	LI	0.162
	UI	0.000
	ADV	0.000

Table 4. Null-subjects acceptance
(in 1st and 2nd persons contexts)

L1	level	p-value
French	LI	0.805
	UI	0.000
	ADV	0.000
Finnish	LI	0.421
	UI	0.000
	ADV	0.000
Arabic	LI	0.049
	UI	0.045
	ADV	NA

Table 5. Null-subjects acceptance
(in 3rd persons contexts)

L1	level	p-value
French	LI	0.222
	UI	0.008
	ADV	0.000
Finnish	LI	0.221
	UI	0.046
	ADV	0.000
Arabic	LI	0.608
	UI	0.128
	ADV	0.030

Table 6. Null-subjects acceptance
(with local antecedents)

L1	level	p-value
French	LI	0.236
	UI	0.090
	ADV	0.000
Finnish	LI	0.721
	UI	0.067
	ADV	0.000
Arabic	LI	0.947
	UI	0.448
	ADV	0.003

Table 7. Null-subjects acceptance
(with non-local antecedents)

L1	level	p-value
French	LI	0.083
	UI	0.000
	ADV	0.000
Finnish	LI	0.212
	UI	0.000
	ADV	0.000
Arabic	LI	0.069
	UI	0.025
	ADV	0.000

Table 8. Missing S-V agreements
acceptance

L1	level	p-value
French	LI	0.535
	UI	0.044
	ADV	0.006
Finnish	LI	0.700
	UI	0.402
	ADV	0.000
Arabic	LI	0.078
	UI	0.348
	ADV	0.043

Table 9. Acceptance of 3rd person singular null-subjects

L1	level	p-value
French	LI	0.412
	UI	0.005
	ADV	0.000
Finnish	LI	0.421
	UI	0.061
	ADV	0.000
Arabic	LI	0.535
	UI	0.235
	ADV	0.023

The Translation task

Table 10. Subject drop

L1	level	p-value
French	LI	0.000
	UI	0.000
	ADV	NA
Finnish	LI	0.001
	UI	0.000
	ADV	NA
Arabic	LI	0.000
	UI	0.000
	ADV	0.000

Table 11. Subject drop (in adverbial clauses)

L1	level	p-value
French	LI	0.000
	UI	0.000
	ADV	NA
Finnish	LI	0.003
	UI	0.000
	ADV	NA
Arabic	LI	0.000
	UI	0.000
	ADV	0.000

Table 12. Subject drop (in complement clauses)

L1	level	p-value
French	LI	NA
	UI	NA
	ADV	NA
Finnish	LI	0.000
	UI	0.000
	ADV	NA
Arabic	LI	0.000
	UI	0.000
	ADV	NA

Table 13. Subject drop (in 3rd person singular contexts)

L1	level	p-value
French	LI	0.000
	UI	0.000
	ADV	NA
Finnish	LI	0.001
	UI	0.000
	ADV	NA
Arabic	LI	0.000
	UI	0.000
	ADV	0.000

Table 14. Missing S-V agreements

L1	level	p-value
French	LI	NA
	UI	NA
	ADV	NA
Finnish	LI	0.000
	UI	0.000
	ADV	NA
Arabic	LI	0.000
	UI	0.000
	ADV	0.000

Appendix 9. Descriptive statistics: acceptance of null subjects with local antecedents vs. with non-local antecedents

Descriptive statistics: acceptance of null subjects with local antecedents vs. with non-local antecedents by French participants

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	Median	mean	StdDev	%
Null subjects with local antecedents	LI	194	100	15	8	10	11.1	2.67	51.5
	UI	688	78	7	0	3	2.52	1.82	11.3
	ADV	644	13	5	0	0	0.48	1.09	2.0
Null subjects with non-local antecedents	LI	33	12	4	0	1	1.33	1.22	36.4
	UI	114	15	2	0	0	0.48	0.63	13.2
	ADV	104	1	1	0	0	0.04	0.19	1.0

Descriptive statistics: acceptance of null subjects with local antecedents vs. with non-local antecedents by Finnish participants

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	Median	mean	StdDev	%
Null subjects with local antecedents	LI	129	63	14	7	11	10.5	2.66	48.8
	UI	389	47	8	0	2	2.76	2.44	12.1
	ADV	1261	31	2	0	0	0.58	0.77	2.5
Null subjects with non-local antecedents	LI	23	7	2	0	1	1.17	0.75	30.4
	UI	61	6	2	0	0	0.35	0.61	9.8
	ADV	211	3	1	0	0	0.06	0.23	1.4

Descriptive statistics: acceptance of null subjects with local antecedents vs. with non-local antecedents by Arabic-speaking participants

Items type	Participant's Level	Included Responses	Total Acceptance	Max	Min	Median	mean	StdDev	%
Null subjects with local antecedents	LI	341	231	21	9	14	14.4	3.08	67.7
	UI	306	80	10	1	6	5.71	2.95	26.1
	ADV	255	11	4	0	0	1.00	1.34	4.3
Null subjects with non-local antecedents	LI	56	38	4	0	3	2.38	1.15	67.9
	UI	55	16	3	0	1	1.14	1.03	29.1
	ADV	44	3	1	0	0	0.27	0.47	6.8

Appendix 10: Proposed descriptive statistics for the English native-speakers control group showing number of dropping subjects

Number of Participants	Included translation	No. of omission	Max	Min	Mean	StdDev	Percentage of omission
7	105*	0	0	0	0.00	0.00	0.0

*included translation = number of the sentences in the task x number of the participants (15x7=105).

Appendix 11. Descriptive statistics: distributions of items with 3rd person singular null subjects and items with missing subject-verb agreement

11-a. Grammaticality judgment task

Descriptive statistics: acceptance distributions of items with 3rd person singular null subject pronouns and items with missing S-V agreement by French participants

Items type	Participant's Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3SG null subjects	LI	140	63	10	2	7.00	2.50	45.0
	UI	487	56	4	0	1.81	1.30	11.5
	ADV	453	8	3	0	0.30	0.67	1.8
Items with missing S-V agreement	LI	79	35	9	0	3.89	3.02	44.3
	UI	292	105	10	0	3.39	2.74	36.0
	ADV	266	63	7	0	2.33	2.17	23.7

Descriptive statistics: acceptance distributions of items with 3rd person singular null subject pronouns and items with missing S-V agreement by Finnish participants

Items type	Participant's Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3SG null subjects	LI	92	47	9	6	7.83	1.17	51.1
	UI	269	34	6	0	2.00	1.84	12.6
	ADV	893	24	3	0	0.45	0.80	2.7
Items with missing S-V agreement	LI	55	21	7	1	3.50	2.07	38.2
	UI	158	43	6	0	2.53	1.46	27.2
	ADV	526	52	5	0	0.98	1.34	9.9

Descriptive statistics: acceptance distributions of items with 3rd person singular null subject pronouns and items with missing S-V agreement by Arabic participants

Items type	Participant's Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3SG null subjects	LI	244	168	15	4	10.5	2.73	68.9
	UI	222	69	9	0	4.93	3.02	31.1
	ADV	183	10	3	0	0.91	1.04	5.5
Items with missing S-V agreement	LI	145	100	10	0	6.25	3.00	69.0
	UI	130	70	9	0	5.00	2.75	53.8
	ADV	108	30	8	0	2.73	2.69	27.8

11-b. The translation task

Descriptive statistics for items produced by French speaking learners: 3rd person subject drop vs. missing subject-verb agreement

Items type	Participant's Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3SG null subjects	LI	81	1	1	0	0.11	0.3	1.2
	UI	278	2	2	0	0.06	0.4	0.7
	ADV	243	0	0	0	0.00	0.0	0.0
Items with missing S-V agreement	LI	54	13	4	0	1.44	1.59	24.1
	UI	185	20	3	0	0.65	1.05	10.8
	ADV	162	7	1	0	0.26	0.45	4.3

Descriptive statistics for items produced by Finnish speaking learners: 3rd person subject drop vs. missing subject-verb agreement

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3SG null subjects	LI	47	5	4	0	0.83	1.6	10.6
	UI	153	3	2	0	0.18	0.5	2.0
	ADV	473	0	0	0	0.00	0.0	0.0
Items with missing S-V agreement	LI	33	7	6	0	1.17	2.40	21.2
	UI	102	9	2	0	0.53	0.72	8.8
	ADV	316	3	1	0	0.06	0.23	0.9

Descriptive statistics for items produced by Arabic speaking learners: 3rd person subject drop vs. missing subject-verb agreement

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3SG null subjects	LI	135	19	6	0	1.19	1.6	14.1
	UI	105	5	3	0	0.36	0.9	4.8
	ADV	99	2	1	0	0.18	0.4	2.0
Items with missing S-V agreement	LI	90	46	6	0	2.88	1.86	51.1
	UI	71	23	5	0	1.64	2.02	32.4
	ADV	66	11	3	0	1.00	1.18	16.7

Appendix 12. Descriptive statistics: distributions of items with 3rd person null subjects and items with 1st or 2nd person null subjects

12-a. Grammaticality judgment task

Descriptive statistics: acceptance distributions of items with 3rd person null subjects and items with 1st or 2nd person null subjects by the French participants

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3 null subjects	LI	178	90	13	6	10.0	2.55	50.6
	UI	628	76	7	0	2.45	1.88	12.1
	ADV	588	12	4	0	0.44	0.93	2.0
Items with 1&2	LI	49	22	5	0	2.44	1.59	44.9
	UI	174	17	2	0	0.55	0.62	9.8
	ADV	160	2	1	0	0.07	0.27	1.3

Descriptive statistics: acceptance distributions of items with 3rd person null subjects and items with 1st or 2nd person null subjects by the Finnish participants

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3 null subjects	LI	118	63	13	8	10.5	2.07	53.4
	UI	352	42	8	0	2.47	2.37	11.9
	ADV	1156	27	3	0	0.51	0.82	2.3
Items with 1&2	LI	34	7	3	0	1.17	1.17	20.6
	UI	98	11	3	0	0.65	0.93	11.2
	ADV	316	7	2	0	0.13	0.39	2.2

Descriptive statistics: acceptance distributions of items with 3rd person null subjects and items with 1st or 2nd person null subjects by the Arabic participants

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3 null subjects	LI	313	216	19	8	13.5	2.97	69.0
	UI	285	80	10	1	5.71	3.38	28.1
	ADV	236	14	4	0	1.27	1.42	5.9
Items with 1&2	LI	84	53	5	1	3.31	1.35	63.1
	UI	76	16	3	0	1.14	0.86	21.1
	ADV	63	0	0	0	0.00	0.00	0.0

12-b. The translation task

Descriptive statistics for items produced by French speaking learners: 3rd person subject drop vs. 1st and 2nd persons subject drop

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3 rd person subject drop	LI	99	1	1	0	0.11	0.3	1.0
	UI	340	3	2	0	0.10	0.4	0.9
	ADV	297	0	0	0	0.00	0.0	0.0
Items with 1 st and 2 nd persons subject drop	LI	36	0	0	0	0.00	0.0	0.0
	UI	124	0	0	0	0.00	0.0	0.0
	ADV	108	0	0	0	0.00	0.0	0.0

Descriptive statistics for items produced by Finnish speaking learners: 3rd person subject drop vs. 1st and 2nd persons subject drop

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3 rd person subject drop	LI	58	5	4	0	0.83	1.6	8.6
	UI	187	3	2	0	0.18	0.5	1.6
	ADV	578	0	0	0	0.00	0.0	0.0
Items with 1 st and 2 nd persons subject drop	LI	21	0	0	0	0.00	0.0	0.0
	UI	67	0	0	0	0.00	0.0	0.0
	ADV	209	0	0	0	0.00	0.0	0.0

Descriptive statistics for items produced by Arabic speaking learners: 3rd person subject drop vs. 1st and 2nd persons subject drop

Items type	Participants Level	Included Responses	Total Acceptance	Max	Min	mean	StdDev	%
Items with 3 rd person subject drop	LI	165	24	8	0	1.50	2.3	14.5
	UI	128	7	4	0	0.50	1.3	5.5
	ADV	121	2	1	0	0.18	0.4	1.7
Items with 1 st and 2 nd persons subject drop	LI	59	4	1	0	0.25	0.4	6.8
	UI	45	2	1	0	0.14	0.4	4.4
	ADV	44	0	0	0	0.00	0.0	0.0

Appendix 13. Further statistical analysis and discussion relevant to footnote number 100

Descriptive statistics: total and subtotal numbers of participants in each group and subgroup who were capable of identifying sentence(s) containing an error, but were unable to correct them.				
Level of Proficiency	Finnish	French	Arabic	Total Number of Participants
Lower-Intermediate	3	1	7	11
Upper-Intermediate	5	4	3	12
Advanced	3	3	0	6
Subtotal	11	8	10	29

Descriptive statistics: total numbers of the rejected sentences with null subjects for unknown reasons (no corrections were provided)				
Level of Proficiency	Finnish	French	Arabic	Total Number of rejected sentences
Lower-Intermediate	8	1	15	24
Upper-Intermediate	13	9	5	27
Advanced	3	5	0	8
Subtotal	24	15	20	59

Based on these statistics, one may conclude that those learners were able to detect the ungrammaticality but were unable to verbalise the error (the missing subject) that rendered the sentences ungrammatical. However, this conclusion cannot be taken for certain as these learners were found to accept null subjects in English. Consider the following table:

Descriptive statistics: total numbers of the accepted items with null subjects for those learners who were unable to verbalise the errors in some of the rejected sentences			
Level of Proficiency	Finnish	French	Arabic
Lower-Intermediate	32	18	103
Upper-Intermediate	24	17	22
Advanced	3	5	0
Subtotal	59	40	125

Another reason that makes one remain sceptical about accepting the conclude that those learners were able to detect the ungrammaticality but were unable to verbalise the error comes from the fact that these learners also were found to reject sentences with null subjects but not based on the error that rendered them ungrammatical. Consider the following table

Descriptive statistics: total numbers of the rejected items with null subjects for wrong reasons (CIF+PIF) only for those learners who were sometimes unable to verbalise the errors in some of the rejected sentences			
Level of Proficiency	Finnish	French	Arabic
Lower-Intermediate	3	10	6
Upper-Intermediate	11	0	9
Advanced	0	0	0
Subtotal	14	10	15

Key: CIF: clearly incorrect – a wrong correction was provided, PIF: possibly incorrect – a wrong correction was provided

Therefore, it can be concluded that null subject is still part of the learners' interlanguage competence even if one assumes that those learners were able to detect the ungrammaticality, but were unable to verbalise the error.

References

- Adger, D. (2003). *Core Syntax: A minimalist Approach*. Oxford, Oxford University Press
- Adolphs, S. and Schmitt, N. (2004). Vocabulary coverage according to spoken discourse context. In P. Bogaards and B. Laufer (eds.) *Vocabulary in a Second Language* (pp. 39-52). Amsterdam: John Benjamins.
- Agustín Llach, M.P. (2011). *Lexical errors and accuracy in foreign language writing*. Channel View Publications, United Kingdom.
- Alderson, J. C., C. Clapham and D. Wall. (1995). *Language test construction and evaluation*. Cambridge, Cambridge University Press.
- Alderson, J.C. (2000). *Assessing Reading*. Cambridge, Cambridge University Press.
- Aleid, S. (2000) The Use of Pictures and Drawings in Teaching English Paragraph Writing in Saudi Arabia Schools, King Saud University: unpublished MA thesis.
- Al-Kasey, T. and A. T. Pérez Leroux. (1998). Second language acquisition of Spanish null subjects. In S. Flynn, G. Martohardjono and W. O'Neil (eds.), *The generative study of second language acquisition* (pp. 161-85). Mahwah, NJ: Lawrence Erlbaum Associates.
- Allen, S. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* 38 (3): 483-521.
- Allen, S. and H. Schroeder. (2003). Preferred argument structure in early Inuktitut spontaneous speech data. In J. Du Bois, L. Kumpf and W. Ashby, (eds.), *Preferred Argument Structure: Grammar as Architecture for Function* (pp. 301-338). Amsterdam: Benjamins.
- Alqurashi, A. (2010). *Free Relative Clauses in Modern Standard Arabic and Hejazi Arabic*. <<http://www.google.co.uk/search?q=Free+Relative+Clauses+in+Modern+Standard+Arabic+and+Hejazi+Arabic&og=Free+Relative+Clauses+in+Modern+Stan>

[dard+Arabic+and+Hejazi+Arabic&sugexp=chrome,mod=16&sourceid=chrome&ie=UTF-8](#)> accessed 17/10/2012.

Al-Seghayar, M. (1997). Remarks on the Types of Complementizers in Arabic. *Essex Graduate Student Papers in Language & Linguistics 1*: 1-12.

http://www.essex.ac.uk/linguistics/publications/egspll/volume_1/index.aspx> accessed 29/10/2012.

Anderson, M. (1992). *Intelligence and Development: A Cognitive Theory*. Oxford, Blackwell.

Anderson, S.R. and D.W. Lightfoot. (2002). *The Language Organ: Linguistics as Cognitive Physiology*. Cambridge University Press, Cambridge.

Aoun, E. J., E. Benmamoun and L. Choueiri. (2010). *The Syntax of Arabic*. Cambridge, Cambridge University Press.

Ariel, M. (1990). *Assessing noun-phrases antecedents*. London, Routledge.

Ariel, M. (1996). Referring expressions and the +/-coreference distinction. In T. Fretheim and J. Gundel (eds.), *Reference and referent accessibility* (pp. 13–35). Amsterdam: John Benjamins.

Aronoff, M. (2003). Null subjects in child language: Evidence for a performance account. In G. Garding and M. Tsujimura (eds.), *WCCFL 22 Proceedings* (pp. 43-55). Somerville, MA: Cascadilla Press.

Ayoun, D. (2005). *Parameter setting in language acquisition*. London, Continuum International Publishing Group.

Azar, B. and S. Hagen. (2011). *Fundamentals of English Grammar*. Pearson, Longman.

Baayen, R. H. (2008). *Analysing Linguistic Data: A practical introduction to Statistics using*. Cambridge, Cambridge University Press.

Bailey, N., C. Madden and S. Krashen. (1974). Is there a 'natural sequence' in adult second language learning? *Language Learning* 24: 235-43.

- Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis, and E. Tognini-Bonelli (eds.), *Text and technology. In honour of John Sinclair*, (pp. 233-250). Philadelphia and Amsterdam: Benjamins.
- Baker, M.C. (1997). Thematic Roles and Syntactic Structure. In L. Haegeman (ed.), *Elements of Grammar. Handbook of Generative Syntax* (pp. 73-137). Kluwer: Dordrecht.
- Bard, E.G., D. Robertson and A. Sorace. (1996). Magnitude estimation of linguistic acceptability. *Language* 72: 32-68.
- Barrios, S. Jiang, N. and Idsardi, W. (2016). Similarity in L2 Phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts. *Second Language Research*, Vol. 32(3) 367–395.
- Beck, M.-L. (1998). L2 acquisition and obligatory head movement: English-speaking learners of German and the local impairment hypothesis. *Studies in Second Language Acquisition* 20: 311-48.
- Bhat, D.N.S. (2007). *Pronouns*. Oxford, Oxford University Press.
- Bialystok, E. and B. Miller. (1999). The problem of age in second language acquisition: influences from language, task, and structure. *Bilingualism: Language and Cognition* 2: 127-45.
- Bianchi, V. (2006). On the syntax of personal arguments. *Lingua* 116(12): 2023–2067.
- Biberauer, T., A. Holmberg., I. Roberts and M. Sheehan. (2010). *Parametric variation: Null Subjects in minimalist theory*. Cambridge, Cambridge University Press.
- Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. New York, Springer.
- Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language* 68: 706-55.

- Birdsong, D. (1999). Introduction: Whys and why nots of the Critical Period Hypothesis for second language acquisition. In D. Birdsong (ed.) *Second Language Acquisition and the Critical Period Hypothesis* (pp. 1-22). Lawrence Erlbaum: Mahwah, NJ.
- Birdsong, D. (2004). Second Language Acquisition and Ultimate Attainment. In A. Davies and C. Elder (eds.), *The handbook of applied linguistics* (pp. 82-105). Oxford: Blackwell Publishing Ltd.
- Birdsong, D. and M. Molis. (2001). On the evidence for maturational effects in second language acquisition. *Journal of Memory and Language* 44: 235-49.
- Bley-Vroman, R. (1989). The logical problem of second language learning. In: S. Gass and J. Schachter (eds.), *Linguistic perspectives on second language acquisition* (pp. 41-68). Cambridge: Cambridge University Press.
- Bley-Vroman, R., S. Felix and G. Ioup. (1988). The accessibility of Universal Grammar in adult language learning. *Second Language Research* 4: 1-32.
- Bloom, L. (1970). *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT Press.
- Bloom, L. (1991) *Language Development from Two to Three*. Cambridge, Cambridge University Press.
- Bloom, L., M. Lahey., L. Hood., K. Lifter and K. Fliess.(1980). Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language* 7: 235–6 [reprinted in Bloom 1991, 261– 289].
- Bloom, L., P. Lightbown and L. Hood. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development* 40 (2): 1-79.
- Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry* 21: 491-504.
- Boeckx, C., M. Horno-Chéliz and J. Mendívil-Giró. (2012). *Language, from a Biological Point of View: Current Issues in Biolinguistics*. Newcastle upon Tyne: Cambridge Scholars Publishing.

- Bolhuis, J. and M. Everaert. (2013). *Birdsong, Speech, and Language: Exploring the Evolution of Mind and Brain*. Cambridge, Mass: MIT Press.
- Borer, H. (1984). *Parametric syntax. Case studies in Semitic and Romance languages*. Dordrecht, Foris.
- Bowerman, M. (1976). *Early Syntactic Development: A Cross-linguistic Study with Special Reference to Finnish*. Cambridge, Cambridge University Press.
- Braidi, S. (1999). *The acquisition of second language syntax*. London, Arnold.
- Bretherton, I., S. McNew and M. Beeghly-Smith. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a 'theory of mind'. In M. E. Lamb and L. R. Sherrod (eds.), *Infant social cognition: empirical and theoretical considerations* (pp. 333-371). Hillsdale, NJ: Erlbaum.
- Bromberg, H. and K. Wexler. (1995). Null Subjects in Questions. In C. T. Schutze, J. B. Ganger and K. Broihier (eds.), *Papers on Language Processing and Acquisition*, MIT Working Papers in Linguistics 26, 221-248.
- Brown, J. D. and T. S. Rodgers. (2002). *Doing Second language Research*. Oxford, Oxford University Press.
- Brown, R. (1973). *A First Language*. Cambridge, MA, Harvard University Press.
- Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research* 6 (2): 145-173.
- Carrol, G. and Conklin, K. (2016). Found in translation: The influence of the L1 on the reading of idioms in a L2. *Studies in Second Language Acquisition* 38: 403–443.
- Carroll, S. and J. Meisel. (1990). Universals and second language acquisition: Some comments on the state of the current theory. *Studies in Second Language Acquisition* 12(2): 201-208.
- Chafe, W. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago, University of Chicago Press.

- Chafe, W. (1996). Inferring identifiability and accessibility. In T. Fretheim and J. Gundel (eds.), *Reference and referent accessibility* (pp. 37–46). Amsterdam: John Benjamins.
- Charney, R. (1980). Speech roles and the development of personal pronouns. *Journal of Child Language* 7: 509–28.
- Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge Mass.
- Chomsky, N. (1972). *Language and Mind* (enlarged edition). Harcourt Brace Jovanovich, New York.
- Chomsky, N. (1975). *Reflections on language*. New York, Random House.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Foris.
- Chomsky, N. (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. MIT Press, Cambridge Mass.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York, Praeger.
- Chomsky, N. (1988). *Language and Problem of Knowledge*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (1989). Some notes on economy of derivation and representation. *MIT Working Papers in Linguistics* 10: 43-74 (reprinted as Chapter 2 of Chomsky 1995).
- Chomsky, N. (1992). On the nature, use, and acquisition of Language. In T. Pütz (ed.), *Thirty Years of Linguistic Evolution: Studies in Honour of René Dirven on the Occasion of His Sixtieth Birthday* (pp. 3-30). John Benjamins Publishing.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass: MIT Press.
- Chomsky, N. (1998). Minimalist inquiries: The framework. *MIT Working Papers in Linguistics* 15: 1–56.

- Chomsky, N. (2000). *New horizons in the study of language and mind*. New York, Cambridge University Press.
- Chomsky, N. (2005). On Phases. In R. Freidin, C. P. Otero and M.-L. Zubizarreta (eds.), *Foundational Issues in Linguistic Theory* (pp. 133-166). MIT Press: Cambridge Mass.
- Christie, K. and J. P. Lantolf. (1992). The ontological status of learner grammaticality judgments in UG approaches to L2 acquisition. *Rassegna Italiana de Linguistica Applicata* 24(3): 31-52.
- Clahsen, H. (1988). Parameterized grammatical theory and language acquisition: a study of the acquisition of verb placement and inflection by children and adults. In S. Flynn and W. O'Neil (eds.), *Linguistic theory in second language acquisition* (pp. 47-75). Dordrecht: Kluwer.
- Clahsen, H. and P. Muysken. (1986). The availability of universal Grammar to adult and child learners: a study of the acquisition of German word order. *Second Language Research* 2: 93-119.
- Clahsen, H. and U. Hong. (1995). Agreement and null subjects in German L2 development: new evidence from reaction-time experiments. *Second language research* 11: 57 – 87.
- Clahsen, H., M. Penke and T. Parodi. (1993/1994). Functional categories in early child German. *Language Acquisition* 3: 395-429.
- Clahsen, H., S. Eisenbeiss and A. Vainikka. (1994). The seeds of structure: a syntactic analysis of the acquisition of Case marking. In T. Hoekstra and B. D. Schwartz (eds.), *Language acquisition studies in generative grammar* (pp. 85-118). Amsterdam: John Benjamins.
- Clahsen, H., S. Eisenbeiss and M. Penke. (1996). Lexical learning in early syntactic development. In H. Clahsen (ed.), *Generative perspectives on language acquisition: empirical findings, theoretical considerations, crosslinguistic comparisons* (pp. 129-59). Amsterdam: John Benjamins.

- Clancy, P. (1993). Preferred argument structure in Korean acquisition. In E. Clark (ed.), *Proceedings of the 25th Annual Child Language Research Forum* (pp. 307-314). Stanford, CA: CSLI.
- Clark, E. (1970). How young children describe events in time. In G. B. Flores D'Arcais and W. J. M. Levelt (eds.), *Advances in Psycholinguistics* (pp. 275–284). Amsterdam: North-Holland.
- Clark, E. (1973). How children describe time and order. In C. A. Ferguson and D. I. Slobin (eds.), *Studies of child language development* (pp. 585–606). New York: Hoil, Rinehart and Winston.
- Clark, E. V. (1978). From gestures to word: On the natural history of deixis in language acquisition. In J. S. Bruner and A. Garton (eds.), *Human growth and development: Wolfson College lectures*. Oxford: Clarendon Press.
- Clough, P. and C. Nutbrown. (2007). *A Student's Guide to Methodology: Justifying Enquiry*, 2nd Edition. London, SAGE Publications.
- Cohen, L., L. Manion and K. Morrison. (2000). *Research Methods in Education*, 5th
- Cohen, L., L. Manion and K. Morrison. (2007). *Research Methods in Education*, 6th Edition. London, Routledge Flamer.
- Cole, M. (2010). Thematic Null Subjects and Accessibility. *Studia Linguistica* 64: 271-320.
- Cook, V. (1994). The metaphor of access to Universal Grammar in L2 learning. In N. Ellis (ed.), *Implicit learning of language* (pp. 477-502). London: Academic Press.
- Cook, V. (1996). *Second language learning and language teaching*. London, Edward Arnold.
- Cook, V. and M. Newson. (1996). *Chomsky's universal grammar*. Oxford, Blackwell Publishers Ltd.
- Cook, V. and M. Newson. (2001). *Chomsky's universal grammar*, 2nd edition. Oxford, Blackwell Publishers Ltd.

- Cook, V. and M. Newson. (2007). *Chomsky's universal grammar*, 3rd edition. Oxford, Blackwell Publishers Ltd.
- Cook, V. (1990). Timed comprehension of binding in advanced L2 learners of English. *Language Learning* 40(4): 544-599.
- Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language* 63: 544-73.
- Corder, S. P. (1967). The significance of learners' errors. *International review of applied linguistics* 5:161-70.
- Corder, S. P. (1973). The elicitation of interlanguage. In j. Svartvik (ed.), *Errata* (pp. 36-47). Stockholm: Rotobekman.
- Costa, J. and P. Lambert. (2009). France and language(s): Old policies and new challenges in education. Towards a renewed framework? In K. Keogh, J. Ní Mhurchú, H. O'Neill and M. Riney (eds.), *Language policy and practice in Europe: Emerging challenges and innovative responses* (pp. 15-26). Brussels: Consortium of Institutions for Development and Research in Education in Europe/Department for Educational Development (CIDREE/DVO).
- Cote, S. A. (1996). Grammatical and Discourse Properties of Null Arguments in English. Unpublished *PhD thesis*, University of Pennsylvania.
<<http://educ.jmu.edu/~cotesa/research.html>> accessed 14/07/2012.
- Cowan, R. and Y.A. Hatasa. (1994). Investigating the validity and reliability of native speaker and second-language learner judgments about sentences. In E. Tarone, S.M. Gass and A. Cohen (eds.), *Research methodology in second language acquisition* (pp.287-302). Hillsdale, NJ: Lawrence Erlbaum.
- Cowart, W. (1997). *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. Thousand Oaks, CA, Sage Publications.
- Creswell, J. W., V. L. Plano Clark, M. L. Gutmann and W. E. Hanson. (2003). Advanced mixed methods research designs. In A. Tashakkori and C. Teddlie (eds.), *Handbook of mixed methods in social and behavioural research* (pp. 209–240). Thousand Oaks, CA: Sage.

- Curtiss, S. (1977). *Genie: A Psycholinguistic Study of a Modern Day "Wild Child"*, Academic Press, London.
- Davidson, B. (1996). Pragmatic 'weight' and Spanish subject pronouns: The pragmatic and discourse uses of 'tú' and 'yo' in spoken Madrid Spanish. *Journal of Pragmatics* 26: 543-565.
- Davies, A. (2008). The Native Speaker in Applied Linguistics. In A. Davies and C. Elder (eds.), *The Handbook of Applied Linguistics*. Oxford: Blackwell.
- Davies, W. D. (1996). Morphological uniformity and the null subject parameter in adult SLA. *Studies in second language acquisition* 18: 475-93.
- DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499-533.
- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. Doughty and M. H. Long (eds.), *The handbook of second language acquisition*, 313–49. Oxford: Blackwell.
- Denscombe, M. (2007). *The good researcher guide: for small-scale social research project*. New York, Open University Press.
- Diessel, H. and M. Tomasello. (2001). The acquisition of finite complement clauses in English: a corpus-based analysis. *Cognitive Linguistics* 12:1–45.
- Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.
- Dimitriadis, A. (1995). When pro-drop languages don't: On overt pronominal subjects in Greek. *Penn Working Papers in Linguistics* 2: 45–60.
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Mahwah, NJ, Lawrence Erlbaum.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford, Oxford University Press.

- du Plessis, J., D. Solin, L. Travis, L. White. (1987). UG or not UG, that is the question: A reply to Clahsen and Muysken. *Second Language Research* 3: 56-75.
- Dulay, H. and M. Burt. (1973). Should we teach children syntax. *Language Learning* 23: 245-58.
- Dulay, H. and M. Burt. (1974). Natural sequences in child second language acquisition. *Language Learning* 24: 37-53.
- Dulay, H., B. Marina and S. D. Krashen. (1982). *Language Two*. Oxford, CUP.
- Dulay, H., B. Marina and S. D. Krashen. (1982). *Language Two*. Oxford, CUP.
- Ellis, N. and A. Beaton. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning* 43: 559-617.
- Ellis, R. (1990). Grammaticality judgements and learner variability. In H. Burmeister and P. L. Rounds (eds.), *Variability in second language acquisition: Proceedings of the Tenth Meeting of the Second Language Research Forum* (pp. 25-60). Eugene, OR: University of Oregon, Department of Linguistics.
- Ellis, R. (1991). Grammaticality judgments and second language acquisition. *Studies in Second Language Acquisition* 13(2): 161-86.
- Epstein, S., S. Flynn and G. Martohardjono. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Brain and Behavioral Sciences* 19: 677-714.
- Epstein, S., S. Flynn and G. Martohardjono. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Brain and Behavioral Sciences* 19: 677-714.
- Eubank, L. (1996). Negation in early German-English interlanguage: More valueless features in the L2 initial state. *Second Language Research* 12: 73-106.
- Eubank, L. And S. Grace. (1998). V-to-I and inflection in non-native grammars. In M-L. Beck (ed.), *Morphology and its interfaces in L2 knowledge* (pp. 69-88). Amsterdam: John Benjamins.

- Fassi Fehri, A. (1993). *Issues in the Structure of Arabic Clauses and Words*. Dordrecht, Kluwer
- Fassi Fehri, A. (2009). Arabic Silent Pronouns, Person, and Voice. *Brill's Annual of Afroasiatic Languages and Linguistics* 1: 3–40.
- Fassi Fehri, A. (2011). Arabic Agree, Silent Pronouns, and Reciprocal. In E. Broselow and H. Ouali (eds.), *Papers from the Annual Symposia on Arabic Linguistics*. John Benjamins.
- Ferguson, C. (1983). Language planning and language change. In H. Cobarrubias and J. Fishman (eds.), *Progress in language planning: international perspectives* (pp. 29–40). Berlin: Mouton.
- Ferreira, F. Bailey, K. G. and Ferraro, V. (2002). Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science*, Vol. 11, pp. 11-15.
- Fitch W.T., M.D. Hauser and N. Chomsky. (2005). The evolution of the language faculty: Clarification and implications. *Cognition* 97: 179-210.
- Flynn, S. (1987). *A Parameter-Setting Model of L2 Acquisition*. Dordrecht, Reidel.
- Flynn, S. (1996). A parameter-setting approach to second language acquisition. In W. Ritchie and T. Bhatia (eds.), *Handbook of Language Acquisition* (pp. 121-58). San Diego: Academic Press.
- Flynn, S., C. Foley and I. Vinnitskaya. (2004). The cumulative-enhancement model for language acquisition: comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *The International Journal of Multilingualism* 1: 3–16.
- Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, Mass: MIT Press.
- Fodor, J.D. and C. Crowther. (2002). Understanding stimulus poverty arguments. *The Linguistic Review* 19: 105-145.
- Frank, M. (1972). *Modern English*. New Jersey, Prentice- Hall, Inc.

- Frascarelli, M. (2007). Subjects, topics, and the interpretation of referential pro. *Natural Language and Linguistic Theory* 25: 691-734.
- Fromkin, V., R. Rodman and N. Hyams. (2013). *An Introduction to Language*. Cengage Learning.
- Gass, S. and A. Beretta. (1991). *Syntactic constraints and the reliability of L2 grammaticality judgments*. Paper presented at AAAL, New York.
- Gass, S. M. (1994). The reliability of second-language grammaticality judgments. In E. E. Tarone, S. M. Gass and A. D. Cohen (eds.), *Research methodology in second language acquisition* (pp. 303–322). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gilquin, G. (2001). The integrated contrastive model: Spicing up your data. *Languages in Contrast* 3 (1): 95-123.
- Gordon, P. (1998). The True-Value judgement task. In D. McDaniel, C. Mckee and H. Cairns (eds.), *Methods for Assessing Children's Syntax*. Cambridge, Mass: MIT Press.
- Goss, N., Z. Ying-Hua and J. P. Lantolf. (1994). Two heads may be better than one: mental activity in L2 grammaticality judgments. In E. E. Tarone, S. M. Gass and A. D. Cohen (eds.), *Research methodology in SLA* (pp. 263-286). Hillsdale, NJ: Lawrence Erlbaum.
- Grami, M.G. (2010). *The Effects of Integrating Peer Feedback into University-Level ESL Writing Curriculum: A Comparative Study in a Saudi Context*. Ph.D. Thesis, Newcastle University.
- Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg and M. Johansson (eds.), *Languages in contrast: Papers from a symposium on text-based cross-linguistic studies, Lund 4-5 March 1994* (pp. 37-51). Lund: Lund University Press.
- Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (ed.), *Learner English on computer* (pp. 3-18). London, UK: Routledge.

- Granger, S. (2003). *The International Corpus of Learner English: A new resource for foreign language learning and teaching and second language acquisition research. TESOL Quarterly* 73 (3): 538-46.
- Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In K. Aijmer (ed.), *Corpora and language teaching* (pp. 13-32). Amsterdam, The Netherlands: John Benjamins.
- Granger, S., Dagneaux, E. and F. Meunier (eds.). (2002). *The international corpus of learner English: Handbook and CD-ROM*. Louvain-la-Neuve, Belgium: Presses Universitaires de Louvain.
- Gray, J. (2000). The ELT Course Book as Cultural Artifact: How Teachers Censor and Adapt. *ELT Journal* 45 (3): 274-283.
- Greenfield, P. and J. H. Smith. (1976). *The structure of communication in early language development*. New York, Academic Press.
- Gries, S. T. (2009). *Statistics for Linguistics with R: A Practical Introduction*. Germaney, Walter de Gruyter.
- Grimshaw, J. and S. Rosen. (1990). Knowledge and Obedience: The Developmental Status of the Binding Theory. *Linguistic inquiry* 21(2): 187–222.
- Guasti, M. T. (2002). *Language acquisition. The growth of grammar*. Cambridge, MA, MIT Press.
- Guerriero, S., A. Cooper, Y. Oshima-Takane and Y. Kuriyama. (2001). A discourse-pragmatic explanation for argument realization and omission in English and Japanese children's speech. In A. Do, L. Domínguez and A. Johansen (eds.), *Proceedings of the 25th Annual Boston University Conference on Language Development* (pp. 319-330). Somerville, MA: Cascadilla Press.
- Gundel, J. K. (1999). On different kinds of focus. In P. Bosch and R. van der Sandt (eds.), *Focus: Linguistic, cognitive, and computational perspectives* (pp. 293-305). Cambridge: Cambridge University Press.

- Gundel, J. K., N. Hedberg and R. Zacharski. (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69: 274-307.
- Gürkanlı, O., M. Nakupoglu and A. Özyürek. (2007). Shared information and argument omission in Turkish. *Proceedings of the 31st Annual BUCLD 2*: 262-273. Somerville, MA: Cascadilla Press.
- Haegeman, L. (1990). Subject pronouns and subject clitics in West Flemish. *The linguistic Review* 7: 333-63.
- Haegeman, L. (1994). *Introduction to government and binding theory* (2nd ed.). Oxford, Blackwell.
- Haegeman, L. (1997). Register variation, truncation, and subject omission in English and in French. *English Language and Linguistics* 1: 233-270.
- Haegeman, L. (2000). Adult null subjects in non-pro-drop languages. In M. -A. Friedemann and L. Rizzi (eds.), *The acquisition of syntax: studies in comparative developmental linguistics* (pp. 129-69). London: Longman.
- Haegeman, L. (2010). Internal syntax of adverbial clauses. In K. Grohmann and I.-M. Tsimpli (eds.), *Exploring the left periphery* (628-648). Thematic issue: *Lingua* 120.
- Haegeman, L. and T. Ihsane. (2001). Adult null subjects in the non-pro drop languages: two diary dialects. *Language Acquisition* 9: 329-346.
- Hamann, C. and K. Plunkett. (1998). Subjectless sentences in child Danish. *Cognition* 69: 35-72.
- Hamann, C. L. Rizzi and U. Frauenfelder. (1996). The acquisition of subject and object clitics in French. In H. Clahsen (ed). *Generative Perspectives on Acquisition* (pp. 309-334). Amsterdam: Benjamins.
- Han, Y. (2000). Grammaticality judgment tests: How reliable are they? *Applied Language Learning* 11(1): 177–204.
- Han, Z-H. (2004). *Fossilization in adult second language acquisition*. Buffalo, NY: Multilingual Matters Limited.

- Harmer, J. (1991). *The practice of English language teaching*. Harlow, England: Longman.
- Hauser, M.D., W.T. Fitch and N. Chomsky. (2002). The faculty of language: what is it, who has it, and How did it evolve? *Science* 298: 1569-1579.
- Hawkins, H., R. Towell and M.N. Lamy. (2001). *French Grammar and Usage*. London, Arnold publishers.
- Hawkins, R. (1998). The inaccessibility of formal features of functional categories in second language acquisition. Paper presented at the Pacific Second Language Research Forum, Tokyo.
- Hawkins, R. (2001a). *Second language syntax: A generative introduction*. Oxford, Blackwell Publishing.
- Hawkins, R. (2001b). The theoretical significance of Universal Grammar in second language acquisition. *Second Language Research* 17: 345-67.
- Hawkins, R. (2005). Revisiting *wh*-movement: the availability of an uninterpretable [*wh*]-feature in interlanguage grammars. In L. Dekydtspotter, R.A. Sprouse and A. Liljestr and (eds), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference (GASLA2004)* (pp. 124–37). Somerville, MA: Cascadilla Proceedings Project.
- Hawkins, R. and C. Chan. (1997). The partial availability of Universal Grammar in second language acquisition: the ‘failed functional features hypothesis’. *Second Language Research* 13: 187-226.
- Hawkins, R. and H. Hattori. (2006). Interpretation of multiple *wh*-questions by Japanese speakers: a missing uninterpretable account. *Second Language Research* 22: 269-301.
- Hawkins, R., R. Towell and N. Bazergui. (1993). Universal grammar and the acquisition of French verb movement. *Second language Research* 9: 189-233.
- Haznedar, B. and Gavruseva, E. (2008). *Current Trends in Child Second Language Acquisition*. Amsterdam: John Benjamins.

- Hedgcock, J. (1993). Well-formed vs. ill-formed strings in L2 metalingual tasks: specifying features of grammaticality judgements. *Second Language Research* 1: 1-21.
- Hegarty, M. (2005). *A Feature-Based Syntax of Functional Categories. The Structure, Acquisition and Specific Impairment of Functional Systems*. Berlin/New York: Mouton de Gruyter.
- Heltai, P. (1992). Lexical Contrasts in Learners' Translations. In K. Klaudy, J. Lambert and A. Sohár. (eds.), *Translation Studies in Hungary* (pp. 71-82). Budapest : Scholastica
- Hilles, S. (1986). Interlanguage and the pro-drop parameter. *Second Language Research* 2: 33-52.
- Hilles, S. (1991). Access to universal grammar in second language acquisition. In L. Eubank (ed.), *Point counterpoint. Universal Grammar in thesecond language* (305-338). Amsterdam/Philadelphia: Benjamins.
- Hirakawa, M. (1993). Null subjects versus null objects in an early grammar of Japanese. *McGill Working Papers in Linguistics* 9: 30-45.
- Holes, C. (2004). *Modern Arabic: Structures, Functions and Varieties*. Washington, D.C., Georgetown University Press.
- Holmberg, A. (2001). The syntax of yes and no in Finnish. *Studia Linguistica* 55: 141-174.
- Holmberg, A. (2005). Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36: 533-564.
- Holmberg, A. (2010a). Null subject parameters. In: T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan (eds.), *Parametric Variation: null subjects in minimalist theory* (pp: 88-124). Cambridge: Cambridge University Press.
- Holmberg, A. (2010b). The null generic subject pronoun in Finnish: a case of incorporation in T. In T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan (eds.), *Parametric Variation: null subjects in minimalist theory* (pp. 200-230). Cambridge: Cambridge University Press.

- Holmberg, A. and M. Sheehan. (2010). Control into finite clauses in partial null-subject languages. In T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan (eds.), *Parametric Variation: Null Subjects in Minimalist Theory* (pp.125-152). Cambridge: Cambridge University Press.
- Holmberg, A., A. Nayudu and M. Sheehan. (2009). Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish, and Marathi. *Studia Linguistica* 63 1: 59-97.
- Hood, L. And L. Bloom. (1979). What, when, and how about why: a longitudinal study of early expressions of causality. *Monographs of the Society for Research in Child Development* 44 (6): 1-44.
- Hopkins, K. D., J.C. Stanley and B. R. Hopins. (1990). *Educational and psychological measurement and evaluation* (7th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Huang, C. T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic inquiry* 15: 531-74.
- Huang, C. T. J. (1989). Pro-drop in Chinese: A generalized control theory. In O. Jaeggli and K. Safir (eds.), *The null subject parameter* (pp. 1-44). London: Springer.
- Huang, Y. (2000). *Anaphora: A cross-linguistic approach*. Oxford, Oxford University Press.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge, Cambridge University Press.
- Hughes, M. and S. Allen. (2006). A discourse-pragmatic analysis of subject omission in child English. *Proceedings of the 30th Annual BUCLD 1*. Somerville, MA: Cascadilla Press, 293-304.
- Hughes, M. and S. Allen. (2008). Child-directed speech and the development of referential choice in child English. Poster presented at the Eleventh Congress of the International Association for the Study of Child Language (IASCL), Edinburgh, Scotland.

- Hulstijn, J. H. (2008). Incidental and Intentional Learning. In C. Doughty and M. Long (eds.), *The handbook of second language acquisition*. London: Wiley-Blackwell.
- Hulstijn, J.H. (2010). Measuring second language proficiency. In E. Blom and S. Unsworth (eds.), *Experimental methods in language acquisition research* (Language learning and language teaching, 27) (pp. 185-200). Amsterdam: Benjamins.
- Hurford, J. (1991). The evolution of the critical period for language acquisition. *Cognition* 40: 159-201.
- Hyams, N. (1986). *Language acquisition & the theory of parameters*. Dordrecht, Reidel.
- Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht, Kluwer.
- Hyams, N. (1992). The genesis of clausal structure. In J. Meisel (ed.), *The Acquisition of Verb Placement; Functional Categories and V2 Phenomena in Language Development* (pp. 371- 400). Dordrecht: Kluwer.
- Hyams, N. (2011). Missing Subjects in Early Child Language. In J. De Villiers and T. Roeper (eds.), *Handbook of Generative Approach to Language Acquisition* (pp. 13-52). Springer: Dordrecht.
- Hyltenstam, K. and N. Abrahamsson.(2003). Maturation constraints in SLA. In C. Doughty and M. Long (eds.), *Handbook of second language acquisition* (pp. 539-588). Oxford: Blackwell.
- Ingham, R. (1998). Tense without agreement in early clause structures. *Language Acquisition* 7: 51-81.
- Ionin, T. (2013). Morphosyntax. In J. Herschensohn and M. Young-Scholten (eds.), *The Cambridge Handbook of Second Language Acquisition* (pp. 505-528). Cambridge: Cambridge University Press.
- Jaeggli, O. (1982). *Topics in Romance Syntax*. Dordrecht, Foris.

- Jaeggli, O. and K. Safir. (1989). The null subject parameter and parametric theory. In O. Jaeggli and K. Safir (eds.), *The null subject parameter* (pp: 1-40). London: Springer.
- Jaensch, C. (2008). L3 acquisition of articles in German by native Japanese speakers. In R. Slabakova, J. Rothman, P. Kempchinsky and E. Gavrusseva (eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)* (pp. 81- 89). Somerville, MA: Cascadilla Proceedings Project.
- Jakubowicz, C., N. Milller, B. Riemer and Rigaut, C. (1997). The case of subject and object omissions in French and German. In E. Hughes, M. Hughes and A. Greenhill (eds.), *Proceedings of the 2^{1st} annual Boston University conference on language development* (pp.331-42). Somerville, Mass: Cascadilla Press.
- Johns, C. (2007). *Interpreting Agreement*. Unpublished *doctoral dissertation*, Durham University.
- Johnson, J.S. and E.L. Newport. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60–99.
- Johnson, J.S., K.D. Shenkman, E. L. Newport and D. L. Medin. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language* 35: 335–52.
- Judy, T. (2011). L1/L2 parametric directionality matters: More on the null subject parameter in L2 acquisition. *EUROSLA Yearbook* 11: 165-90.
- Judy, T. and J. Rothman. (2010). From a superset to a subset grammar and the Semantic Compensation Hypothesis: Subject pronoun anaphora resolution evidence in L2 English. In K. Franich, K.M. Iserman and L.L. Keil (eds.) *BUCLD 34: Proceedings of the 34th annual Boston University Conference on Language Development* (pp. 197-208). Somerville, MA: Cascadilla Press.
- Källkvist, M. (1998). How different are the results of translation tasks? A study of lexical errors. In K. Malmkjær (ed.), *Translation and Language Teaching: Language Teaching and Translation*. Manchester: st. Jerome Publishing.

- Kanno, K. (1997). The acquisition of null and overt pronominals in Japanese by English speakers. *Second Language Research* 13: 265-87.
- Kanno, K. (1998). The stability of UG principles in second language acquisition. *Linguistics* 36: 1125-46.
- Keller, F. (1999). Grammaticality Judgments and Linguistic Methodology. *Journal of Logic, Language and Information* 8(1): 114-121.
- Kempson, R. (1996). Semantics, pragmatics and interpretation. In S. Lappin (ed.), *The handbook of contemporary semantic theory* (pp. 561–598). Oxford: Blackwell.
- Kenstowicz, M. (1989). The null subject parameter in modern Arabic dialects. In O. Jaeggli and K. Safir (eds.). *The null subject parameter*. London: Springer.
- Kidd, E., E. Lieven and M. Tomasello. (2005). The acquisition of complement clause constructions: A sentence repetition study. *Proceedings of the 32nd Stanford Child Language Forum* (pp. 50-59).
- Kinn, K. (2013). Null arguments in Old Norwegian in an early Germanic perspective. Paper presented at the 21st International Conference on Historical Linguistics, Oslo, Norway.
- Klein-Braley, C. and V. Smith. (1985). Incalculable and full of risks?: Translation L1 to L2 as a testing procedure". In C. Titford and A. E. Hieke (eds.), *Translation in Foreign Language Teaching and Testing*. Tübingen: Narr.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.
- Lado, R. (1961). *Language Testing*. New York NY, McGraw Hill.
- Lafond, L. (2001). The pro-drop parameter in second language acquisition revisited: a developmental account. Unpublished PhD thesis, University of South Carolina.
- Lakshmanan, U. (1994). *Universal grammar in child second language acquisition*. Amsterdam, Benjamins.

- Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge, Cambridge University Press.
- Lantolf, J. P. (1990). Reassessing the null-subject parameter in second language acquisition. In H. Burmeister and P. L. Rounds (eds.), *Variability in second language acquisition: Proceedings of the Tenth Meeting of the Second Language Research Forum* (pp. 929-452). Eugene, OR: University of Oregon, Department of Linguistics.
- Lantolf, J.P. and W. Frawley. (1988). Proficiency: Understanding the construct. *Studies in Second Language Acquisition* 10: 181-195.
- Lardiere, D. (1998a). Case and tense in the 'fossilized' steady state. *Second Language Research* 14: 1-26.
- Lardiere, D. (1998b). Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research* 14: 359-75.
- Lardiere, D. (2000). Mapping features to forms in second language acquisition. In J. Archibald (ed.), *Second language acquisition and linguistic theory*. Oxford: Blackwell, 102–29.
- Lardiere, D. (2007). Acquiring (or assembling) functional categories in second language acquisition. In A. Belikova, L. Meroni and M. Umeda (eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, 233-244.
- Lardiere, D. (2008). Feature-assembly in second language acquisition. In J. Liceras, H. Goodluck and H. Zobl (eds.), *The role of formal features in second language acquisition* (pp. 107–40). New York: Lawrence Erlbaum.
- Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25 (2): 173-227.
- Larsen-Freeman, D. and M. Long. (1991). *An introduction to second language acquisition research*. London, Longman.

- Lasnik, H. and J. Uriagereka. (2002). On the poverty of the challenge. *The Linguistic Review* 19: 147-150.
- Léautaud, P. (1988). Le Fléau. *Journal Particulier* . 1917- 39. Paris: Mercure de France.
- Lee, C. W. (2014). Learning and unlearning object drop in anaphoric and non-anaphoric contexts in L2 English. Unpublished PhD thesis, University of Newcastle upon Tyne.
- Lee, S.-Y. (2008). Argument-adjunct asymmetry in the acquisition of inversion in *wh* questions by Korean learners of English. *Language Learning*, 58, 625-663.
- Lenneberg, E. (1967). *Biological Foundations of Language*. Wiley, New York.
- Leonardi, V. (2010). *The role of pedagogical translation in second language acquisition: from theory to practice*. Switzerland, Peter Lang AG.
- Leung, Y.-K.I. (2006). Full transfer vs. partial transfer in L2 and L3 acquisition. In R. Slabakova, S. Montrul and P. Prévost (eds.), *Inquiries in linguistic development: in honor of Lydia White* (pp. 157–87). Amsterdam: John Benjamins.
- Leung, Y.-K.I. (2007). L3 acquisition: why it is interesting to generative linguists. *Second Language Research* 23: 95–114.
- Levy, Y., and A. Vainikka. (1999). The development of a mixed null subject system: A cross linguistic perspective with data on the acquisition of Hebrew. *Language Acquisition* 8: 363-84.
- Liceras, J. (1989). On some properties of the “pro-drop” parameter: Looking for missing subjects in non-native Spanish. In S.Gass and J. Schachter (eds.), *Linguistic Perspectives on Second language Acquisition* (pp. 109-133). Cambridge: Cambridge University Press.
- Liceras, J. M. and L. D’az. (1999). Topic-drop versus pro-drop: Null subjects and pronominal subjects in the Spanish L2 of Chinese, English, French, German, Japanese and Korean speakers. *Second Language Research* 15(1): 1-40.
- Liceras, J., L. Díaz and D. Maxwell. (1999). Null subjects in non-native grammars: the Spanish L2 of Chinese, English, French, German, Japanese and Korean speakers.

- In E. C. Klein and G. Martohardjono (eds.), *The development of second language grammars. A generative approach* (pp. 109-146). Amsterdam/Philadelphia: Benjamins.
- Lightbown, P. (1987). Classroom language as input to second language acquisition. In C. W. Pfaff (ed.), *First and Second Language Acquisition Processes*. Rowley, MA: Newbury House.
- Lightfoot, D. (1999). *The Development of Language: Acquisition, Change and Evolution*. Blackwell, Oxford.
- Limber, J. (1973). The genesis of complex sentences. In T. E. Moore (ed.), *Cognitive development and the acquisition of language* (pp. 169 – 85). New York, NY: Academic Press.
- Long, M. (1997). *Fossilization: Rigor mortis in living linguistic systems?* Plenary speech at EUROSLA 7, Barcelona.
- Long, M. (2003). Stabilization and fossilization in interlanguage development. In C. Doughty and M. Long (eds.), *Handbook of second language acquisition* (pp. 487-536). Oxford: Blackwell.
- Low, G. (1999). What respondents do with questionnaires: Accounting for incongruity and fluidity. *Applied Linguistics 20*: 503-33.
- Lowie, W. and M. H. Verspoor. (2015). Variability and Variation in Second Language Acquisition Orders: A Dynamic Reevaluation. *Language Learning 65*: 63-88.
- Mackey, A. and S. Gass. (2005). *Second language Research: Methodology and Design*. London, Routledge.
- Makino, T. (1993). *Perspectives on Second Language Acquisition*. Tokyo, Yumi Press.
- Malmkjær, K. (1998). Introduction: Translation and Language Teaching. In K. Malmkjær (ed.), *Translation and Language Teaching: Language Teaching and Translation*. Manchester: st. Jerome Publishing.
- Margaza, P. and A. Bel. (2006). Null Subjects at the Syntax-Pragmatics Interface: Evidence from Spanish Interlanguage of Greek Speakers. In M. G. O'Brien, C.

- Shea and J. Archibald (eds.) *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp.88-97). Somerville, MA: Cascadilla Press.
- Martinez, E., and C. Godev. (1994). Should reading comprehension be tested in the target or the native language? A pilot Study. *ERIC Document Reproduction*
- Mazuka, R., B. Lust, T. Wakayama and W. Snyder. (1986). Distinguishing effects of parameters in early syntax acquisition: A cross-linguistic study of Japanese and English. *Papers and Reports on Child Language Development* 25:73-82.
- McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics* 21: 395-423.
- Meisel, J. (1991). Principles of Universal Grammar and strategies of language learning: some similarities and differences between first and second language acquisition. In L. Eubank (ed.), *Point counterpoint: Universal grammar in the second language* (pp. 231-276). Amsterdam: John Benjamins.
- Meisel, J. (1997). The acquisition of the syntax of negation in French and German: contrasting first and second language acquisition. *Second Language Research* 13: 227-63.
- Messick, S. (1989). Validity. In R. L. Linn (ed.), *Educational measurement* (3rd ed., (pp. 13–103). New York, NY: Macmillan.
- Mitchell, R. and F. Myles. (2004). *Second language learning theories*. London, Arnold.
- Montalbetti, M. (1984). After Binding: On the interpretation of pronouns. Unpublished doctoral dissertation, MIT.
- Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: a case of morphosyntactic convergence. *Bilingualism: Language and Cognition* 7 (2): 125–42.
- Montrul, S., R. Foote and S. Perpifian. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: the effects of age and context of acquisition. *Language Learning* 58: 502-53.

- Murphy, R. (2010). *English Grammar in Use: A Self-study Reference and Practice Book for Intermediate Students of English*. Cambridge, Cambridge University Press.
- Murphy, V. (1997). The effect of modality on a grammaticality judgement task. *Second Language Research* 13 (1): 34-65.
- Neeleman, A. and F. Weerman. (1997). LI and L2 word order acquisition. *Language Acquisition* 6: 125-70.
- Newson, D. (1998). Translation and Foreign Language Teaching. In K. Malmkjær (ed.), *Translation and Language Teaching* (pp. 63-68). Manchester: St. Jerome Publishing.
- O'Grady, W. (1997) *Syntactic Development*. University of Chicago Press.
- Orfitelli, R. and N. Hyams. (2008). An experimental study of children's comprehension of null subjects: Implications for grammatical/performance accounts. *Proceedings of the 32nd Annual BUCLD 2*: 335-346. Somerville, MA: Cascadilla Press.
- Orfitelli, R., and T. Grüter. (2013). Do null subjects really transfer? In J. Cabrelli-Amano, T. Judy and D. Pascual y Cabo (eds.), *Proceedings of the 12th Generative Approaches to Second Language Acquisition Conference (GASLA 2013)* (pp.145- 54). Somerville, MA: Cascadilla Press.
- Orsolini, M. and P. Di Giacinto. (1996). Use of referential expressions in 4-year-old children's narratives: invented versus recalled stories. In C. Pontecorvo, M. Orsolini, B. Burge and L. Resnick (eds.), *Children's early text construction* (pp. 67–81). Mahwah, NJ: Erlbaum.
- Orsolini, M., F. Rossi and C. Pontecorvo. (1996). Re-introduction of referents in Italian children's narratives. *Journal of Child Language* 23: 465–486.
- Ortega, L. and G. Iberri-Shea. (2005). Longitudinal research in SLA: Recent trends and future directions. *Annual Review of Applied Linguistics* 25: 26–45.
- Ortega, L. and H. Byrnes. (eds.). (2008). *The longitudinal study of advanced L2 capacities*. New York: Routledge.

- Osterhout, L., A. Kim and G. Kuperberg. (2012). The Neurobiology of Sentence Comprehension. In M. Spivey, M. Joannisse and K. McRae (eds.), *The Cambridge Handbook of Psycholinguistics* (pp. 365-389). Cambridge: Cambridge University Press
- Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning* 30: 440-72.
- Paradis, M. (2004). A neurolinguistic theory of bilingualism. Amsterdam: John Benjamins.
- Payne, M. and M. Almansour. (2014). Foreign language planning in Saudi Arabia: beyond English. *Current Issues in Language Planning* 15(3): 327-42.
- Perez, J. R. (2006). Teaching English in the Expanding Circle. Germany, GRIN Verlag.
- Pérez-Leroux, A. T. and W. Glass. (1997). OPC effects on the L2 acquisition of Spanish. In A. T. Pérez-Leroux and W. Glass (eds.), *Contemporary perspectives on the acquisition of Spanish. Vol. 1: Developing Grammar* (pp. 149-65). Somerville, MA: Cascadilla Press.
- Perez-Lérroux, A.T. and W. Glass. (1999). Null anaphora in Spanish second language acquisition: probabilistic versus generative approaches. *Second Language Research* 15: 220-49.
- Perlmutter, D. (1971). *Deep and surface constraints in syntax*. New York: Holt, Rinehart and Winston.
- Pery-Woodley, M.P. (1990). Contrasting discourses: contrastive analysis and a
- Pesetsky, D. and E. Torrego. (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language* (pp. 355-426). MIT Press: Cambridge Mass.
- Pesetsky, D. and E. Torrego. (2007). The syntax of valuation and the interpretability of features. In S. Karimi, V. Samiiian and W. Wilkins (eds.), *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation* (pp. 262-94). John Benjamins: Amsterdam.

- Phinney, M. (1987). The pro-drop parameter in second language acquisition. In: T. Roeper and E. Williams (eds.), *Parameter setting* (pp. 221-238). Berlin and New York: Springer.
- Pierce, A. (1992). *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars*. Dordrecht, Kluwer.
- Pinker, S. (1995). Language Acquisition. In: L. Gleitman and M. Liberman (eds.), *An innovation to cognitive science* (pp. 107-134). Cambridge, Mass: MIT Press.
- Pollitt, A. (no date) *The Oxford Online Placement Test: The Meaning of OOPT Scores*.
<http://www.oxfordenglishtesting.com/uploadedFiles/Buy_tests/oopt_meaning.pdf> accessed 25/06/2014.
- Pollock, J.-Y.(1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365-424.
- Prévost, P. (2009) *The Acquisition of French : The development of inflectional morphology and syntax in L1 acquisition, bilingualism, and L2 acquisition*. John Benjamins Publishing, Amsterdam/Philadelphia.
- Prévost, P. and L. White. (2000). Missing Surface Inflection or Impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16: 103- 33.
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning* 52: 513–36.
- Quesada, M. L and S. E. Blackwell. (2009). The L2 Acquisition of Null and Overt Spanish Subject Pronouns: A Pragmatic Approach. In J. Collentine, M. García, B. Lafford and F. Marín (eds.), *Proceedings of the 11th Hispanic Linguistics Symposium* (pp. 117-30). Somerville, MA: Cascadilla Press.
- Radford, A. (1986). Small children's small clauses. *Research papers in linguistics* 1: 1-38. University College of North Wales: Bangor.

- Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax*. Cambridge, MA: Blackwell.
- Radford, A. (2004). *Minimalist Syntax: Exploring the Structure of English*. Cambridge, Cambridge University Press.
- Rahman, M. and E. Alhaisoni. (2013). Teaching English in Saudi Arabia: Prospects and challenges. *Academic Research International Journal* 4 (1): 112-18.
- Rasetti, L. (2000). Null subjects and root infinitives in child grammar of French. In M. Friedemann and L. Rizzi (eds.), *The Acquisition of Syntax* (pp.236-68). Harlow, England: Longman.
- Redman, S. (2003). *English Vocabulary in Use: Pre-intermediate and Intermediate*. Cambridge, Cambridge University Press.
- Richards, N. (2001). *Movement in Language: Interactions and Architectures*. Oxford, Oxford University Press.
- Richards, J. C., J. Platt and H. Platt. (1992). *Longman Dictionary of Language Teaching and Applied Linguistics*. London, Longman.
- Ringbom, H. (2002). The status of English in Finland. In S. Granger, E. Dagneaux and F. Meunier (eds.), *The international corpus of learner English* (pp. 93-96). Louvain-la-Neuve, Belgium: Presses Universitaires de Louvain.
- Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht, Foris.
- Rizzi, L. (1986). On the status of subject clitics in Romance. In O. Jaeggli and C. Silva-Corvalan (eds), *Studies in Romance linguistics* (pp. 391-420). Dordrecht: Foris.
- Rizzi, L. (1992). Early null subjects and root null subject. In T. Hoekstra and B. Schwartz (eds.), *Language acquisition studies in generative grammar* (pp. 151-76). Amsterdam: Benjamins.
- Rizzi, L. (1994). Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition* 3: 371-93.

- Rizzi, L. (1998). Remarks on Early Null Subjects. In A. Greenhill, M. Hughes, H. Littlefield and H. Walsh (eds), *Proceedings of the 22nd Annual Boston University Conference on Language Development* (pp. 14-38). Somerville, MA: Cascadilla Press.
- Rizzi, L. (2000). Remarks on early null subjects. In M. Friedemann and L. Rizzi (eds.), *The Acquisition of Syntax* (pp. 269-92). Harlow, England: Longman.
- Rizzi, L. (2005a). On the grammatical basis of language development: A case study. In G. Cinque and R. Kayne. (eds.), *The Oxford Handbook of Comparative Syntax* (pp. 70-109). New York: Oxford University Press.
- Rizzi, L. (2005b). Grammatically-based target-inconsistencies in child language. In K. Deen, J. Nomura, B. Schulz and B. Schwartz. (eds.), *Proceedings of the Inaugural Conference of GALANA* (pp. 19-49). Cambridge, MA: UCONN/MIT Working Papers in Linguistics.
- Robert, I. and A. Holmberg. (2005). On the role of parameters in Universal Grammar: a reply to Newmeyer. In H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz and J. Koster (eds.), *Organizing Grammar: Linguistic studies in Honor of Henk Van Riemsdijk* (pp.538-553.). Berlin: Mouton de Gruyter.
- Roberts, I. (2007). *Diachronic Syntax*. New York, Oxford University Press.
- Roberts, I. and A. Holmberg. (2010). Parameters in minimalist theory. In T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan (eds.), *Parametric Variation: null subjects in minimalist theory* (pp. 1-57). Cambridge: Cambridge University Press
- Rodrigues, C. (2000). Deriving referential null subjects from movement. Ph.D. dissertation, University of Maryland.
- Roeper, T. and B. Rohrbacher. (2000). Null Subjects in Early Child Language and the Economy of Projection. In S. Powers and C. Haaman (eds.), *The Acquisition of Scrambling and Cliticization* (pp.345-396). Dordrecht: Kluwer.
- Rom, A. and R. Dgani. (1985). Acquiring case-marked pronouns in Hebrew: the interaction of linguistic factors. *Journal of Child Language* 12: 61–77.

- Rothman, J. (2008). How Pragmatically Odd! Interface Delays and Pronominal Subject Distribution in the L2 Spanish of English Natives. *Studies in Hispanic and Lusophone Linguistics* 1(2): 317-39.
- Rothman, J. (2009). Pragmatic deficits with syntactic consequences?: L2 pronominal subjects and the syntax-pragmatics interface. *Journal of Pragmatics* 41 (5): 951-73.
- Rothman, J. and J. Cabrelli Amaro. (2010). What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research* 26(2): 189–218.
- Rothman, J. and M. Iverson. (2007a). On L2 Clustering and Resetting the Null Subject Parameter in L2 Spanish: Implications and Observations. *Hispania* 90(2): 329–342.
- Rothman, J. and M. Iverson. (2007b). Input Type and Parameter Resetting: Is Naturalistic Input Necessary? *International Review of Applied Linguistics* 45(4): 285–319.
- Rowlett, P. (2007). *The Syntax of French*. Cambridge, Cambridge University Press.
- Sampson, G. (2002). Exploring the richness of the stimulus. *The Linguistic Review* 19: 73-104.
- Sano T. and N. Hyams. (1994). Agreement, finiteness, and the development of null arguments. *Proceedings of NELS* 24: 543-58.
- Sauter, K. (2002). Transfer and access to Universal Grammar in adult second language acquisition. Ph.D. dissertation, University of Groningen.
<<http://dissertations.ub.rug.nl/FILES/faculties/arts/2002/k.sauter/thesis.pdf>>
accessed 28/05/2012.
- Schachter, J. (1989). Testing a proposed universal. In S. Gass and J. Schachter (eds.), *Linguistic perspectives on second language acquisition* (pp. 73-88). Cambridge: Cambridge University Press.
- Schachter, J. (1990). On the issue of completeness in second language acquisition. *Second Language Research* 6: 93-124.

- Schachter, J., and V. Yip. (1990). Grammaticality judgments: why does anyone object to subject extraction? *Studies in Second Language Acquisition* 12(4): 379-92.
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schmitt, N., X. Jiang and W. Grabe. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal* 95(1): 26–43.
- Scholz, B.C. and G. K. Pullum. (2002). Searching for arguments to support linguistic nativism. *The Linguistic Review* 19: 185-223.
- Schoonen, R. (2011). How language ability is assessed. In E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning*, Volume II. (pp. 701-716). New York: Routledge.
- Schütze, C. (2011). Linguistic evidence and grammatical theory. *Wiley Interdisciplinary Reviews: Cognitive Science* 2(2): 206-21.
- Schütze, C. T. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistics methodology*. Chicago, The University of Chicago.
- Schwartz, B and R. Sprouse. (2013). Generative approaches and the poverty of stimulus. In J. Herschensohn and M. Young-Scholten (eds.), *The Cambridge Handbook of Second Language Acquisition* (pp. 137-158). Cambridge: Cambridge University Press.
- Schwartz, B. and R. Sprouse. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research* 12: 40–72.
- Schwartz, B., and R. Sprouse. (1994). Word order and nominative case in non-native language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra and B. Schwartz (eds.), *Language acquisition studies in generative grammar: Papers in Honour of Kenneth Wexler from the 1991 GLOW Workshops* (pp. 317-68). Amsterdam: John Benjamins.

- Sciullo, A. and S. Boeckx. (2011). *The Bilingual Enterprise New Perspectives on the Evolution and Nature of the Human Language Faculty*. Oxford, Oxford University Press.
- Seliger, H. and E. Shohamy. (1989). *Second language research methods*. Oxford, Oxford University Press.
- Selinker, L. (1972). Interlanguage. *International review of applied linguistics* 10: 209-31.
- Selinker, L. (1989). CA/EA/IL: The earliest experimental record. *International Review of Applied Linguistics* 27: 267 - 91.
- Selinker, L. (1992). *Rediscovering Interlanguage*. London: Longman.
- Selinker, L. (1996). On the Notion of 'IL Competence' in Early SLA Research: An Aid to Understanding Some Baffling Current Issues. In G. Brown, K. Malmkjær and J. N. Williams (eds.), *Performance and Competence in Second Language Acquisition* (pp. 89-113). Cambridge: Cambridge University Press.
- Serratrice, L. (2005). The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics* 26: 437-462.
- Serratrice, L. and A. Sorace. (2003). Overt and Null Subjects in Monolingual and Bilingual Italian Acquisition. *Proceedings of the 27th Annual BUCLD 2*: 739-750. Somerville, MA: Cascadilla Press.
- Shlonsky, U. (2009). Hebrew as a partial null-subject language. *Studia Linguistica*, 63: 133–57.
- Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing* 1(2): 147-70.
- Sigurðsson, H. Á. (2004). The syntax of Person, Tense, and speech features. *Italian Journal of Linguistics* 16: 219–51.
- Silva, M. (1991). Simultaneity in children's narratives: the case of *when*, *while* and *as*, *Journal of Child Language* 18: 641–662.

- Skehan, P. (1989). *Individual differences in second language learning*. London, Edward Arnold.
- Slabakova, R. (2009). What is easy and what is hard to acquire in a second language. *Proceedings of the GASLA 10*: 280-94.
- Smith, N. (2004) *Chomsky: Ideas and Ideals*, 2nd edition. Cambridge, Cambridge University Press.
- Smith, N.V. and I. M. Tsimpli. (1995). *The mind of a savant: language learning and modularity*. Oxford, Blackwell.
- Solon, M. (2016). Do learners lighten up? Phonetic and Allophonic Acquisition of Spanish // by English-Speaking Learners. *Studies in Second Language Acquisition 1*: 1-32.
- Sorace A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: data, interpretations and methods. *Bilingualism: Language and Cognition 7* (2): 143–45.
- Sorace, A. (1993). Incomplete and divergent representations of unaccusativity in nonnative grammars of Italian. *Second Language Research 9*: 22-48.
- Sorace, A. (1996). The use of acceptability judgments in second language acquisition research. In W. C. Ritchie and T. K. Bhatia (eds.), *Handbook of second language acquisition* (pp. 375–409). San Diego, CA: Academic.
- Sorace, A. (2003). Optimality as a feature of L2 end-state grammars. In C. Doughty and M. Long (eds.), *Handbook of second language acquisition* (pp. 130-152). Oxford: Blackwell.
- Spada, N., and P. Lightbown. (1999). Instruction, first language influence and developmental readiness in second language acquisition. *The Modern Language Journal 83*: 1-22.
- Stibbard, R. M. (1998). The principled use of translation in foreign language teaching. In K. Malmkjaer (ed.), *Translation and language teaching: Language teaching and translation* (pp. 69-79). Manchester, England: st.Jeromse.

- Sudman, S. and N. M. Bradburn. (1983). *Asking Questions*. San Francisco, CA: Jossey-Bass.
- Sutter, J.C. and C.J. Johnson. (1990). School-age children's metalinguistic awareness of grammaticality in verb form. *Journal of Speech and Hearing Research* 33: 84-95.
- Syed, Z. (2003). TESOL in the gulf: The sociocultural context of English language teaching in the gulf. *TESOL Quarterly*, 37(2), 337-341
- Tarone, E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning* 35: 373-404.
- Tarone, E. (1988). *Variation in interlanguage*. London, Edward Arnold.
- Terrazas Gallego, M. and M.P. Agustin Llach. (2009). Exploring the increase of receptive vocabulary knowledge in the foreign language: A longitudinal study. *International Journal of English Studies* 9 (1): 113-33.
- Thomas, M. (1994). Assessment of L2 proficiency in second language acquisition research. *Language Learning* 44: 307-36.
- Thomas, M. (2002). Development of the concept of “the poverty of stimulus”. *The Linguistic Review* 19: 51-71.
- Thomas, M. (2004). *Universal Grammar in Second-Language Acquisition: A History*. London, Routledge.
- Towell, R. (2013). Learning mechanisms and automatization. In J. Herschensohn, and M. Young-scholten (eds), *The Cambridge handbook of second language acquisition*, 114-136. Cambridge: Cambridge University Press.
- Towell, R. and R. Hawkins. (1991). *Approaches to second language acquisition*. Clevedon, Multilingual matters.
- Tremblay, A. (2005). Theoretical and methodological perspectives on the use of grammaticality judgment tasks in linguistic theory. *Second Language Studies* 24: 129-167.

- Tsimplici, I.-M. and Roussou, A (1991). Parameter resetting in L2. *UCL Working Papers in Linguistics* 3: 149-69.
- Tsimplici, I.-M, A. Sorace, C. Heycock and F. Filiaci. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism* 8: 257–77.
- Tsimplici, I-M. (2003). Clitics and Determiners in L2 Greek. Proceedings of Generative Approaches to Second Language Acquisition. Somerville, MA: Cascadilla Press
- Tsimplici, I-M., and M. Dimitrakopoulou. (2007). The interpretability hypothesis: evidence from wh-interrogatives in second language acquisition. *Second Language Research* 23: 215-42.
- Ullbarri, D. M., M. L. Spencer and G. A. Rivas. (1981). Language proficiency and academic achievement: A study of language proficiency tests and their relationship to school ratings as predictors of academic achievement. *Journal of the National Association for Bilingual Education* 5(3): 47-80.
- Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: the declarative/procedural model. In C. Sanz (ed.), *Mind and context in adult second language acquisition: methods, theory and practice*, 141–78. Washington, DC: Georgetown University Press.
- Vainikka, A, and Y. Levy. (1999). Empty subjects in Finnish and Hebrew. *Natural Language and Linguistic Theory* 17: 613-71.
- Vainikka, A. and M. Young-Scholten. (1994). Direct access to X' theory: Evidence from Korean and Turkish adults learning German. *Second Language Research* 12: 7-39.
- Vainikka, A. and M. Young-Scholten. (1996a). The early stages of adult L2 syntax: additional evidence from Romance speakers. *Second Language Research*, 15: 140-76.
- Vainikka, A. and M. Young-Scholten.(1996b). Gradual development of L2 phrase structure. *Second Language Research* 12: 7-39.

- Vainikka, A. and M. Young-Scholten (2011). *The acquisition of German: Introducing the Organic Grammar*. Berlin: De Gruyter Mouton.
- Valian, V. (1990). A problem for parameter-setting models of language acquisition. *Cognition* 35 (2): 105-22.
- Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition* 40: 21–81.
- Valian, V. and Z. Eisenberg. (1996). Syntactic subjects in the spontaneous speech of Portuguese-speaking children. *Journal of Child Language* 23(1): 103-128.
- Valian, V., J. Hoeffner and S. Aubry. (1996). Young children's imitation of sentence subjects: Evidence of processing limitations. *Developmental Psychology* 32 (1): 153-64.
- Vecchio, A. and M. Guerrero. (1995). *Handbook of English Language Proficiency Tests*. New Mexico, New Mexico Highlands University.
<http://www.ncela.gwu.edu/files/rcd/BE020503/Handbook_of_English.pdf>
accessed 24/07/13.
- Verspoor, M. H., W. Lowie and M. van Dijk. (2008). Variability in second language development from a dynamic systems perspective. *Modern Language Journal* 92: 214–231.
- Walkden, G. (2014). *Syntactic Reconstruction and Proto-Germanic*. Oxford, Oxford University Press
- Wang, Q., D. Lillo-Martin, C. Best and A. Levitt. (1992). Null subject versus null object: Some evidence from the acquisition of Chinese and English. *Language Acquisition* 2: 221-54.
- Weir, C. J. (2005). *Language Testing and Validation: An Evidence-Based Approach*. New York, Palgrave MacMillan.
- Wexler, K. (2015). *Do any of our Models of Parameter-Setting Work? Just-so Stories aren't Sufficient*. York University Workshop.
<https://www.york.ac.uk/media/languageandlinguistics/documents/conferences/wlvcc2015/Wexler_handout.pdf> accessed 19/06/2015.

- White, L. (1985). The pro-drop parameter in adult second language acquisition. *Language Learning* 35: 47-62.
- White, L. (1986). Implications of parametric variation for adult second language acquisition. An investigation of the pro-drop parameter. In V. Cook (ed.), *Experimental approaches to second language acquisition* (pp. 55-72). Oxford: Pergamon Press.
- White, L. (1989). *Universal Grammar and second language acquisition*. Amsterdam: John Benjamins.
- White, L. (1992). Long and short verb movement in second language acquisition. *Canadian Journal of Linguistics* 37: 273-86.
- White, L. (1998). *Universal Grammar in second language: the nature of interlanguage representation*.
<www.inflrc.hawaii.edu/networks/nw09/white.pdf>accessed 02/06/2009.
- White, L. (2000). Second language acquisition: from initial to final state. In J. Archibald (ed.), *Second language acquisition and linguistic theory* (pp. 130-55). Oxford: Blackwell.
- White, L. (2003). *Second language acquisition and universal grammar*. Cambridge, Cambridge University Press.
- White, L. (2008). On the Nature of Interlanguage Representation: Universal Grammar in the Second Language. In C. Doughty and M. Long (eds), *The handbook of second language acquisition* (pp. 19-42). Oxford: Blackwell Publishers Ltd.
- White, L. (2010). Second language acquisition at the interfaces. *Lingua* 121: 577-90.
- White, L. and E. Genesee. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research* 11: 233-65.
- Whitfield, B. and J. Pollard. (1998). Awareness Raising in the Saudi Arabian Classroom. In J. C. Richards (ed.), *Teaching in Action* (pp. 143 – 49). Alexandria: TESOL.

- Williams, S. and B. Hammarberg. (1998). Language switches in L3 production: implications for a polyglot speaking model. *Applied Linguistics* 19: 295–333.
- Wistner, B., H. Sakai and M. Abe. (2009). An analysis of the Oxford placement test and the Michigan English placement test as L2 proficiency tests (pp. 33-44). Bulletin of the Faculty of Letters, Hosei University.
- Wright, C. (2010). I still can't questions: the role of working memory in the longitudinal development of L2 English questions in an immersion setting. Unpublished PhD thesis, University of Newcastle upon Tyne.
- Yuan, B. (1997). Asymmetry of null subjects and null objects in Chinese speakers' L2 English. *Studies in second language acquisition* 19: 467–97.
- Zobl, H. (1992). Prior linguistic knowledge and the conservatism of the learning procedure: grammaticality judgments of unilingual and multilingual learners. In S. Gass and L. Selinker (eds.), *Language transfer in language learning, revised edition* (pp. 177–96). Amsterdam: John Benjamins.
- Zyzik, E. (2008). Null objects in second language acquisition: grammatical vs. performance models. *Second language research* 24: 65-110.